Universidade de Aveiro Departamento de Educação e Psicologia
Ano 2018

PEDRO JOÃO BEM-HAJA GABRIEL FERREIRA

PSICOFISIOLOGIA DO TESTEMUNHO OCULAR

PSYCHOPHYSIOLOGY OF EYEWITNESS TESTIMONY

Universidade de Aveiro Departamento de Educação e Psicologia
Ano 2018

PEDRO JOÃO BEM-HAJA GABRIEL FERREIRA

PSICOFISIOLOGIA DO TESTEMUNHO OCULAR

PSYCHOPHYSIOLOGY OF EYEWITNESS TESTIMONY

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Psicologia, realizada sob a orientação científica da Doutora Isabel Maria Barbas dos Santos, Professora Auxiliar do Departamento de Educação e Psicologia da Universidade de Aveiro e do Doutor Carlos Fernandes da Silva, Professor Catedrático do Departamento de Educação e Psicologia da Universidade de Aveiro.

Apoio financeiro da FCT e do POPH/FSE no âmbito do III Quadro Comunitário de Apoio (Refª. da Bolsa: SFRH/BD/85928/2012).

If Schrödinger’s cat can live in two different boxes being the same cat, you and I can also live in two bodies, being one.

To my quantum superposition, Vânia.

o júri

Presidente	Professora Doutora Anabela Botelho Veloso Professora Catedrática, Universidade de Aveiro
	Prof.ª Doutora Sandra Cristina Oliveira Soares Professora Auxiliar, Universidade de Aveiro
	Prof. Paulo Joaquim Fonseca Da Silva Farinha Rodrigues Professor Auxiliar, Universidade da Beira Interior
	Prof.ª Sara Margarida Soares Ramos Fernandes Professora Auxiliar, Universidade Portucalense Infante D. Henrique
	Prof.ª Doutora Joana Patrícia Pereira Carvalho Professora Auxiliar, Universidade Lusófona
	Prof.ª Doutora Isabel Maria Barbas dos Santos Professora Auxiliar, Universidade de Aveiro

agradecimentos

To all those who, challenging Newton's Law of Universal Gravitation, made it possible for me to lift off from the ground throughout this journey

palavras-chave

Testemunho ocular, alinhamentos policiais, psicofisiologia, potenciais evocados, electromiografia facial, ritmo cardíaco, rastreamento ocular.

resumo

As testemunhas oculares são muitas vezes o único meio que temos para aceder à autoria de um crime. Contudo, apesar dos 100 anos de evidência de erros no testemunho ocular, a consciência das suas limitações como meio de prova só ganhou força no advento do ADN. De facto os estudos de exoneração mostraram que 70 % das ilibações estavam associadas a erros de testemunho ocular. Estes erros têm um impacto social elevado principalmente os falsos positivos, por colocar inocentes na prisão. De acordo com a literatura, deverão ser utilizadas novas abordagens para tentar reduzir o numero de erros de identificação. Destas abordagens, destacam-se a análise dos padrões de movimentos oculares e os potenciais evocados. Nos nossos estudos utilizamos essas novas abordagens com o objetivo de examinar os padrões de acerto ou de identificação do criminoso, usando um paradigma de deteção de sinal.

No que diz respeito aos movimentos oculares, não foram encontrados padrões robustos de acerto. No entanto, obtiveram-se evidências oculométricas de que a fusão de dois procedimentos (Alinhamento Simultâneo depois de um Alinhamento Sequencial com Regra de Paragem) aumenta a probabilidade de acerto. Em relação aos potenciais evocados, a P100 registou maior amplitude quando identificamos um inocente. Este efeito é concomitante com uma hiperactivação no córtex prefrontal ventromedial (CPFVM) identificada na análise de estimação de fontes. Esta hiperativação poderá estar relacionada com uma exacerbação emocional da informação proveniente da amígdala. A literatura relaciona a hiperativação no CPFVM com as falsas memorias, e estes resultados sugerem que a P100 poderá ser um promissor indicador de falsos positivos. Os resultados da N170 não nos permitem associar este componente ao acerto na identificação. Relativamente à P300, os resultados mostram uma maior amplitude deste componente quando identificamos corretamente um alvo, mas não diferiu significativamente de quando identificamos um inocente. Porém, a estimação de fontes mostrou que nessa janela temporal (300-600 ms) se verifica uma hipoativação dos Campos Oculares Frontais (COF) quando um distrator é identificado. Baixas ativações dos COF estão relacionadas com redução da eficiência de processamento e com a incapacidade para detetar alvos. Nas medidas periféricas, a eletromiografia facial mostrou que a maior ativação do corrugador e a menor ativação do zigomático são um bom indicador de quando estamos perante um criminoso. No que diz respeito ao ritmo cardíaco, a desaceleração esperada para os alvos devido à sua saliência emocional apenas foi obtida quando a visualização de um alvo foi acompanhada por um erro na identificação (i.e., um falso negativo). Neste trabalho de investigação parece que o sistema nervoso periférico está a responder corretamente, identificando o alvo, por ser emocionalmente mais saliente, enquanto que a modulação executiva efectuada pelo CPFVM conduz ao falso positivo. Os resultados obtidos são promissores e relevantes, principalmente quando o resultado de um erro poderá ser uma condenação indevida e, consequentemente, uma vida injustamente destruída.

keywords

Eyewitness, lineups, psychophysiology, event-related potentials, facial electromiography, heart rate, eye-tracking.

abstract

Eyewitnesses are often the only way we can access the author of a crime. However, despite 100 years of evidence of errors in eyewitness testimony, awareness of its limitations only gained strength with the advent of DNA. In fact, 70% of exonerations have been associated with eyewitness errors. These errors have a high social impact, mainly false positives. According to the literature, new approaches to try to reduce the number of identification errors should be used. Of these, the study of oculometric patterns and event-related Potentials (ERP) stand out. In our studies, these new approaches were used with the objective of examining patterns of accuracy, using a signal detection paradigm. Regarding eye movements, no entirely clear patterns were found. However, there was oculometric evidence that the merging of two procedures (Simultaneous Lineup after a Sequential Lineup with Stopping Rule) increases performance accuracy. Regarding ERPs, the P100 registered a larger amplitude when an innocent was identified. This effect is concomitant with a hyperactivation in the ventromedial prefrontal cortex (VMPFC) identified by source estimation analysis. This hyperactivation might be related to an emotional exacerbation of the information coming from the amygdala. The literature relates the hyperactivation in the VMPFC with false memories, and these results suggest that the P100 component might be a promising marker of false positive errors. The results of the N170 do not allow to associate this component with accuracy. Regarding the P300, the results showed a greater amplitude of this component when a target was correctly identified but did not differ significantly from when an innocent was identified. However, source analysis in this time window (300-600 ms) showed a hypoactivation of Frontal Eye Fields (FEF) when a distractor was identified. FEF inactivations are related to the reduction of processing efficiency and to the inability to detect a target. Concerning the peripheral measures, facial electromyography showed that the greater activation of the corrugator and the lower activation of the zygomaticus are a good marker of when we are facing a perpetrator. Regarding heart rate, the expected deceleration for the targets due to their emotional salience was only obtained when the visualization of a target was accompanied by an error in the identification (i.e., a miss). In this research it seems that the peripheral nervous system is responding correctly, identifying the target, because it is emotionally more salient, while the executive modulation carried out by the VMPFC causes the false positive error. The results presently obtained are promising and relevant, especially when the result of an error might be an undue condemnation of an innocent and consequently a destroyed life.

Table of Contents

List of Figures

List of Tables

List of abbreviations

SU	Show-up
ROC	Receiver Operating Characteristic
IGI	Information Gathering Interview
ERP	Event-Related Potentials
EEG	Electroencephalography
BOLD	Blood-oxygen-level dependent
SIML	Simultaneous Lineup
SEQL	Sequential Lineup
SEQL-StopR	Sequential with Stopping rule
SEQL-Passive	Sequential Lineup without behavioural response
ROI	Regions of Interest
CRV	Continuous Random Variable
ID	Identified Individual non-target
EOG	Electrooculogram
EMG	Facial Electromyography
HR	Heart Rate
ECG	Electrocardiography
FEF	Frontal Eye Fields
MNI	Montreal Neurological Institute (Brain Coordinates)

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

INTRODUCTION

> **INTRODUCTION**
>
> There are about fifty psychological laboratories in the United States alone. The average educated man has hitherto not noticed this. If he chances to hear of such places, he fancies that they serve for mental healing, or telepathic mysteries, or spiritistic performances. What else can a laboratory have to do with the mind? Has not the soul been for two thousand years the domain of the philosopher? What has psychology to do with electric batteries and intricate machines? Too often have I read such questions in the faces of visiting friends who came to the Harvard Psychological Laboratory in Emerson Hall and found, with surprise, twenty-seven rooms overspun with electric wires and filled with chronoscopes and kymographs and tachistoscopes and ergographs, and a mechanic busy at his work.

(Hugo Münsterberg, 1908, p. 3)

I believe that the above excerpt (dated 1908) could have been written yesterday and find it rather unfortunate that, 100 years later, psychology is still seen by many as a non-scientific subject, like ‘magic arts’. However, I referenced it because it is part of the first-ever published document that addressed/systematized the psychological basis of errors in eyewitness testimony (Münsterberg, 1908). In this work, called “On the Witness Stand: Essays on Psychology & Crime,” Münsterberg discusses topics such as the illusions of memory in a criminal context, the fallibility of the eyewitness memory or the false confessions, etc. He complemented this with experimental studies that, although not applied to the forensic setting, shared experimental paradigms and, thus, were generalizable. The following year (1909), Whipple published his study “The observer as a reporter: A survey of the 'psychology of testimony'”, in Psychological

Bulletin (Whipple, 1909), where he argued that variables such as age, pathological conditions, time interval, repetition or type of questions may influence the accuracy of reports or testimony. After these two founding studies, several court cases were marked-out for having 'inaccurate identifications' (Massen v. State, 1969; People v. Chambers, 1969; People v. Stanton, 1969; State v. Burch, 1969; State v. Parker, 1969; Stovall v. Denno, 1967). However, it was only in 1974, in an experimental social study conducted by Buckhout, Alper, Chern, Silverberg, and Slomovits that the first objective results of eyewitness performance were obtained. In a sample of 52 undergraduates of Brooklyn College, Buckhout et al. found 13.5% of positive identifications, 13.5% impeached (impeached their identification with another choice), 40.3% mistaken identifications, and 19.2% of nonidentifications. In the same year (1974), Buchkout decided to publish an article in *Scientific American*, affirming "eyewitness testimony is unreliable" (p. 23). However, the reaction of the criminal system was strongly adverse, mainly because it would never be possible to eliminate eyewitness testimony altogether from forensic practices (Wells, 1978). In 1978, Hastie, Landsman, & Loftus, claimed that this aversion to using scientific psychology findings in court was due to the scarcity of publications on relevant laboratory and field research on the subject. In fact, until the middle 70's, only a few studies made significant progress to the eyewitness literature and its applicability (Wells, Lindsay, & Ferguson, 1979). So much so that Garry Wells, amongst other eyewitness researchers, stressed the importance of applying the knowledge of experimental psychology to the judicial world (Wells, 1978). Wells and his colleagues (1979) expressed this concern in "Guidelines for empirically assessing the fairness of a lineup", where they systematically reviewed variables such as the number of distractors, as well as similarities between the target and the distractors in the lineup, which are important factors for a fair lineup. Indeed, the "lineup" procedure is one of the most relevant "system variables" (Wells, 1978).

Procedure Battles: From the Show-ups to the "Sequential vs. Simultaneous" Duel

Lineups were established in the mid-19th century in England in response to the "show-up", which was considered unfair to the suspect (Devlin, 1976). The "show-up" (SU) is an identification procedure in which a single suspect is shown to a witness/victim. However, the suggestiveness of this procedure has been well documented (see Steblay, 2006). An example of this is the study carried out by Yarmey, Yarmey, and Yarmey, (1996), who found that the number of false positives (identification of an innocent) was significantly higher in SUs than in lineups with six members. Many years after replacing this procedure, studies that use the signal detection paradigm and the Receiver Operating Characteristic (ROC) curves for analysis, have confirmed that the show-ups promote more false positive errors, especially in children (Key et al., 2015; Lindsay, Pozzulo, Craig, Lee, & Corber, 1997). Moreover, according to Wetmore et al. (2015), an SU never resulted in better identification accuracy than a lineup. The lineup is an identification procedure in which a group of people (usually six) is presented to an eyewitness to identify or exclude the suspect(s) (Devlin, 1976). Lineups are used to clarify the participation of a particular suspect in a crime (Demarchi, 2013) and are expected to ensure an identification (or the absence of one) considered fair, mainly for those who are innocent (Steblay, 2006). If the suspect does not match the witness's memory, the probability of being considered guilty decreases significantly (Demarchi, 2013). This lineup is used since the beginning and, according to the Police Executive Research Forum (Agencies, 2013), the 'simultaneous lineup' is still the most commonly used in the United States. In this type of lineup, the members (usually six people - one suspect and five fillers) are presented simultaneously (Wogalter, Malpass, & Mcquiston, 2004).

Despite the superiority of the simultaneous lineup over the show-up, experimental psychology has shown that this procedure may implicate a relative judgment, which promotes the positive identification of somebody, whether it is the suspect or a distractor (Demarchi, 2013). Indeed, as early as 1984, Garry Wells showed that a witness only makes an absolute judgment when they identify the member of the lineup who best matches his memory of the perpetrator. Therefore, in a simultaneous lineup, the eyewitness tends to compare all the lineup members and pick the person who most resembles the suspect - frequently leading to incorrect identifications (false positives/False Alarms).

In order to take advantage of the benefits of the lineups, yet trying to reduce the false positive identifications provoked by a relative judgment, Lindsay and Wells (1985) have proposed the Sequential Lineup where suspects are presented sequentially, promoting an absolute judgment since the witness compares each of the individuals with his or her memory and not with the other members of the lineup[1]. The use of the "stopping rule" is critical in this procedure, for it implies that, as soon as a positive

[1] To answer the social critique that "Psychology is not useful to society," Scott O. Lilienfeld in his article "*Public skepticism of psychology: Why many people perceive the study of human behavior as unscientific*", published in American Psychologist in 2012, listed five great discoveries of psychology useful to society. One of these five great discoveries reported by the author is the Sequential Lineup, placed at the same level as the Kahneman's discovery that gave him the Nobel Prize for Economics.

identification is made, the showing of the pictures must stop, allowing certainty in the identification and preventing the relative comparison - by not being able to see the following nor to go back (Lindsay & Wells, 1985). The sequential lineup procedure is the most commonly used procedure in Canada (Beaudry & Lindsay, 2006). In co-witness protocols (when two or more witnesses viewed the crime), the sequential lineup proved to be the best method, regardless of whether there was any discussion between witnesses about the crime before identification (Yarmey & Morris, 1998). According to Meissner, Tredoux, Parker, and MacLin (2005), the absolute judgment present in the sequential lineup is consistent with the conservative criteria of signal detection theory, resulting in a significantly smaller number of incorrect identifications. Due to this absolute judgment, the change from target to distractor in picking should be majorly reduced in sequential lineups, relatively to simultaneous ones (Clark & Davey, 2005). The superiority of the sequential lineup was also postulated by Steblay, Dysart, Fulero, and Lindsay (2001), who showed that the identification of innocents when the target was present was significantly greater in the simultaneous lineup than in the sequential type. In the same study, the authors also verified a significantly higher number of Correct Rejections for the sequential lineup. More, Lindsay and Bellinger (1999) found that a simultaneous lineup with photographs leads to an increase in the number of false positive identifications. The simultaneous lineup will inspire the witness to use deduction, increasing the tendency to make a false positive error, while a sequential lineup only allows recall (Penrod & Bornstein, 2007). However, if it is true that the sequential lineup provides fewer false positives, it also provides a lower rate of culprit identifications (e.g., Levi, 2016; Mansour & Flowe, 2010; Mecklenburg, 2006). Memon and Gabbert (2003) showed that the sequential lineup reduced both the number of false and correct identifications, among young and older adults, when the target was present. The same pattern was found by Rose, Bull, and Vrij (2005). In order to check pros and cons of each lineup procedure and to verify the lineup with better overall results, in a meta-analysis conducted by Steblay et al. (2011; weighing 72 studies, 13143 participants) the supremacy of the sequential lineup (referring to correct identifications) was demonstrated. This predominance was primarily explained by the fact that the use of absolute judgment decreases the number of false positives (Pozzulo, Reed, Pettalia, & Dempsey, 2016; Steblay, Dysart, Fulero, & Lindsay, 2001; Steblay, Dysart, & Wells, 2011). Nevertheless, according to Gronlund, Wixted, and Mickes (2014), the best lineup is the one that capitalizes the capacity of witnesses to discriminate between guilty and innocent (discriminability). Considering this postulate, Mickes, Flowe, and Wixted (2012) using ROC curves showed that the sequential lineup seems to be poorer that simultaneous in discriminability (the degree to which eyewitnesses can tell the difference between innocent and guilty suspects). The higher discriminability of Simultaneous Lineup is indicated by a ROC curve that bows farther up and away from the 'chance performance' (diagonal line). The same pattern of results was verified by Carlson and Carlson (2014) and by Wixted and Mickes (2014).

This disagreement of views and results regarding the superiority of each type of lineup shows that it is not yet unblemished which one will prove to be undoubtedly superior (Gronlund, Wixted, & Mickes, 2014). Faced with this dichotomy there appears to be only consensus in the fact that the sequential lineup, by promoting an absolute judgment and a conservative behavior, allows the obtaining of fewer false positives (Steblay et al., 2001; Steblay, Dietrich, Ryan, Raczynski, & James, 2011; Wells, Dysart, & Steblay, 2015; Wells, Memon, & Penrod, 2006), and the simultaneous lineup, by promoting a relative judgment and a liberal behavior, allows more Hits in the criminals (Gronlund et al., 2014; Lindsay, Mansour, Beaudry, Leach, & Bertrand, 2009a; Malpass, 2006; Wells, 2014).

Malpass (2006) introduced some apparently inconclusive moral and political notions into the debate: should we increase the probability of identifying the perpetrator at the risk of raising the possibility of a mistaken identification? Or should mistaken identifications be avoided at all costs, enabling a criminal to escape sentence? Wells (2014, p. 14) advances with the same question "Should we adopt a new procedure (simultaneous) that increases the chances that the guilty might be identified, but also increases the chances of mistaken identification?".

Inasmuch as none of the types of lineup demonstrated to be superior, Demarchi (2013) argues that we must use both methods complementarily and, take into account the characteristics of the investigation, we must choose the appropriate type of lineup. If the suspect is arrested shortly after the crime and there is very little (if any) incriminating evidence (which implies a significant margin of error), the sequential lineup is preferable. On the other hand, if the prosecution has a considerable amount of incriminating evidence, the simultaneous lineup is advisable. In the same sense, a crime with a more extensive criminal frame (number of years in reclusion) should be advised the use of sequential lineup to minimize the possibility of convicting an innocent.

Scientific Evidence as the Answer

Given the uncertainty of which type of lineup to use, psychological science has a lot to say about the correct and fair lineup construction. Numerous experimental studies that used lineups have helped to build a number of guidelines to conduct a scientific and ethically fair lineup, always with the aim of decreasing the number of incorrect identifications (e.g., Malpass & Tredoux, 2008; Mcllister, Michel, Tarcza, & Fitzmorris, 2008; Wells & Bradfield, 1999; Wells, Leippe, & Ostrom, 1979).

On the basis of the experimental method, an analogy can be used to compare the construction of a lineup with experimenting. In the literature, it is called the "analogy of Wells" (Wells & Luus, 1990). First, the police have a hypothesis (suspected of a culprit) and assemble the necessary materials to test it (photos of the suspect and distractors). Subsequently, the police create the experimental design (for example, placing the suspect's image in a specific position), instruct participants (eyewitnesses), perform the procedure (show the lineup), record the data (identification or not) and finally, through the collected data, interpret the hypothesis (Santos et al., 2014).

This analogy between a lineup and an experiment helps to clarify the factors that should be taken into account when building a lineup. Some of these factors are listed below.

Similarity between target and distractors

The first step to build a scientific and fair lineup is to create a procedure where the potential suspect is placed amongst plausible distractors. However, according to Corey, Malpass, and McQuiston (1999), low-similarity lineups are often used in real criminal investigations. Thus, the distractors must be selected on the basis of the witness's description of the offender (Wells et al., 1998) and all the members of the lineup must match that description (Doob & Kirshenbaum, 1973; Wells, Seelau, Rydell, & Luus, 1994). In fact, as early as 1980, Lindsay and Wells showed that lineups with high-similarity between members produced fewer identifications of the criminal but also of innocent distractors (i.e., false positives), thus being fairer. Five years later, the same authors have shown that even when telling participants/witnesses that the culprit may not be present in the alignment, the identification of an innocent is still 2.3 times superior when faced with a low-familiarity lineup (Lindsay & Wells, 1985). Similarly, Lindsay, Martin, and Webber (1994) showed that the number of correct identifications increased with the selection of distractors that corresponded precisely to the description of the suspect. Another study showed that the best postdictors of accuracy were the time the witness/participant takes to make a decision and lineup fairness (Smith, Lindsay, & Pryke, 2000). Evidence also shows that for the two types of the lineup (sequential and simultaneous), the fairness of the lineups is the best predictor of accuracy in the identification, regardless of encoding optimality[2] (Tredoux, Parker, & Nunez, 2007).

Multiple witnesses

When the case involves two or more witnesses, an individual lineup is required for each one, ensuring that the distractors match each witness's description (Luus & Wells, 1991). Due to memory bias, the interaction between witnesses of the same event

[2] High encoding optimality - witness has a good memory of the perpetrator.
Levels of encoding optimality: (a) high was a slow-motion 130 s recording, (b) moderate was a normal speed 66 s recording, and (c) low was an edited normal speed 40 s recording

was prohibited in the late 70’s (Warnick & Sanders, 1980a). A study carried out by Levet (2013) showed that witnesses who hear the co-witness pick a person from the lineup are more likely to choose a person from the lineup than witnesses who have not heard any information from the co-witness, or who have heard the co-witness reject the lineup. In the same study, Levett argued that the confidence expressed by a witness is influenced by the trust of the co-witnesses in lineup decision. Levett stated that in cases where there are multiple witnesses, their identifications should not be independent pieces of evidence (Levett, 2013). In a recent study, Rose and Beck (2016), corroborated the fallibility of ocular testimony, mainly when there are multiple witnesses, due to contamination. In spite of this, there is literature that concludes that joining two witnesses in collaborative dyads, where they discuss and make a joint decision, can improve the recall of information and reduce false positives (Yarmey & Morris, 1998).

Position of the suspect in the lineup

In a study conducted by O' Connell and Synnott (2009), an association between position 1 and the low number of correct identifications was obtained. Indeed, the proportion of correct responses was 7.1% for position 1, being significantly lower than positions 3 (50.0%), 4 (64.3%), and 5 (21.4%). As is common practice in experimental psychology studies, randomizing the location of the suspect in the lineup reduces false identifications (Palmer, Sauer, & Holt, 2017). However, this only applies to cases with more than one eyewitness.

Size of the lineup

As we have seen above, one of the oldest topics of discussion in eyewitness is the size of the lineup. The number of lineup members varies significantly across different countries. In the United States, it is usual to have six members in the lineup (e.g., Wells et al., 1998), whereas in the UK and Canada the required minima are 11 and 12 members, respectively (Pozzulo, Dempsey, & Wells, 2010). Despite this disparity, no statistically significant differences were detected between the various sizes of lineups, in adults (e.g., Levi, 2007; Nosworthy & Lindsay, 1990) and children (Pozzulo et al., 2010). Also, Lindsay, Smith, & Pryke (1999) showed that lineup size is rarely associated with the number of false positive identifications. Some authors have pointed out that lineup sizes can be kept to a minimum (even with three distractors alongside the suspect), provided that the high similarity between the suspect and the distractors is maintained (Malpass & Tredoux, 2008; Nosworthy & Lindsay, 1990).

Instructions and interviews

For a long time, the concern to study the impact of the information-gathering interview on the memory of the eyewitness has existed (see Loftus, 1975). Information gathering interview (IGI) refers to the type of police interview, where the interviewer does not confront the suspect with the accusation, encouraging him to talk about what happened, through the use of open questions (e.g., "Describe in as much detail as possible what happened in morning of Christmas day?") (Fisher, McCauley, & Geiselman, 1994; Vrij, Mann, & Fisher, 2006). The IGI is most recommended (see Vrij et al., 2006), opposing the accusatory-type interview, where the interviewer confronts the suspect with an accusation (e.g., "You are hiding something from me") (Henry &

Gudjonsson, 2007). Elizabeth Loftus, in 1975, showed that interviewing immediately after a crime/event can introduce unnecessary information, causing memory reconstruction or distortion (Loftus, 1975). Loftus and Zanni (1975) found the same result and alerted to the consequences that the reconstruction of memory can have in legal cases. Another of the most discussed issues on that topic is that of using misleading questions when interviewing witnesses. In fact, Weinberg, Wadsworth, and Baron (1983) have suggested that misleading questions have a significant impact on the performance of an eyewitness, increasing the possibility of misidentification. In the same way, the results of Hosch, Leippe, Marchioni, and Cooper (1984) showed that the witnesses who had been given biased instructions presented more errors in identifications and rejected fewer lineups.

Warnick and Sanders (1980b) have shown that encouraging the witnesses to say "I don’t know" promotes the reduction of false positives without compromising the proportion of correct responses.

Regarding the type of answers that may be predictors of accuracy, Dunning and Stern (1994) have shown that responses such as "His face just popped out at me" indicate a more automatic and therefore more likely to be the correct response. On the contrary, responses such as "I compared the photos to each other to narrow the choices" are typical of an inaccurate witness, which result from a strategy of elimination.

Despite the absence of written guidelines in most US jurisdictions on how officers must interact with eyewitnesses in lineup procedures (Greene & Evelo, 2014), Wise, Cushman, and Safer (2012) suggest that the lineup administrator should follow the principles below:

1) It is equally important to clear innocents, as it is to identify the criminal;

2) The culprit's appearance may have changed since the crime;

3) The culprit may or may not be in the lineup;

4) The administrator should not know the suspect's identity (double-blind procedure);

5) The investigation should continue independently of whether or not a positive identification is made.

Despite this systematization by Wise et al. (2012), a study carried out by Molinaro, Arndorfer, and Charman (2013) showed that principle 2) (The culprit's appearance may have changed since the crime) could be problematic. In fact, the authors show that the appearance-change instruction inflated false identifications.

Double-blind procedure and presence of police officers

The presence of a police officer at the time of identification is an undeniable reality. However, in addition to the type of questions and the neutral and objective manner in which those questions should be asked, the presence of a lineup administrator who knows who the suspect is, can compromise the validity of the identification (Wells et al., 1998). According to Haw and Fisher (2004), if the police officer responsible for the lineup knows who the suspect is, they can implicitly transmit the identity of the suspect to the witness. Actually, it has been shown that eyewitnesses tend to make the identification decision according to the expectations of the administrator when the level of contact between the two is high (Haw & Fisher, 2004).

A study by Greathouse and Kovera (2009) showed that the diagnostic value of identifications of the culprit in double-blind procedures (where the lineup administrator does not know the identity of the suspect) was twice that of single blinds procedures. In the same study, the authors showed that the administrators exhibit more biased behaviors during single-blind procedures and, when bias factors that are typical of this type of procedure are present, the propensity of the eyewitness to guess increases. According to Dysart, Lawson, and Rainey (2012), the double-blind administration is a prophylactic procedure against the effects of the administrator feedback.

Thus, using a concept that is typical in experimental psychology, the administration of lineups should always obey the double-blind paradigm: the police officer/researcher who conducts the lineup/experiment should not know which member/stimulus of the lineup/task is the suspect/target (Harris & Rosenthal, 1985).

Mugshot

To carry out a lineup, it is imperative to ensure that the witness did not see the suspect in a different situation from the crime scene. This is because when an eyewitness has seen the suspect again, before the lineup (e.g., in a mug-book, police station), there is a higher chance that the witness will identify this suspect in the lineup, even when he/she is innocent (Brown, Deffenbacher, & Sturgill, 1977; Wise et al., 2012). Also, Hilgendorf and Irving (1978) found that witnesses who see the photographs before had a significantly greater proportion of false positives on the following lineup.

According to Gronlund, Carlson, Dailey, and Goodsell (2009), the commitment and the familiarity are the two major explanatory theories for the negative influence of a mugshot on the performance of an eyewitness in the lineup. The first one, as its name implies, is related to the "committing to the previously selected lineup member", i.e., when an eyewitness has picked a particular lineup member from an initial group of mugshots, they are likely to pick that same person again in a later lineup identification (Gorenstein & Ellsworth, 1980). Thus, this “committing to a selection strategy” is related to the fact that mugshot choosers continue to choose in lineups (Memon, Hope, Bartlett, & Bull, 2002).

The familiarity effect can be interpreted by the source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993) or fuzzy-trace theory (Reyna & Brainerd, 1995). The source-monitoring framework postulates that an eyewitness has trouble in discerning between distinct memory traces (e.g., a filler vs. culprit) and might not recognize that a feeling of familiarity during the lineup may be due to other past exposure to that filler rather than having, in fact, witnessed that individual performing the crime (Lindsay, 1994).

The fuzzy-trace theory suggests that previous exposure to a mugshot reinforces the representation of the individual present in the mugshot in the deterrence of the individual seen at the crime scene, making a familiar distractor more likely to be identified (Reyna & Brainerd, 1995).

Although we can control, using scientific evidence, most of these lineup variables that are related to identification errors, other variables such as the characteristics of the witness, the event, or the perpetrator are uncontrollable (Narby, Cutler, & Penrod, 1996). However, if we are aware of the effects of these variables, we can attribute a degree of validity to the eyewitness testimony. Thus, we subsequently will discuss some variables of witnesses, events, and perpetrators that studies of experimental psychology identified as relevant.

Variables of the Eyewitness

Personality

According to Clifford and Scott (1978), there is no relationship between accuracy in identification and personality. However, a study conducted by Megreya and Bindemann (2013) showed that women with high levels of trait anxiety and low levels of emotional regulation give significantly more misidentification errors. Also, in a recent study, Curley, MacLean, and Murray (2017) showed that, of the Five-Factor Model, only openness to experience (dimension) is correlated with the number of false positives and correct responses. That is, people with high levels of openness were more likely to be correct and produced a lower number of false positives.

Age

The age of the eyewitness is one of the best-studied variables in this area. Parker and Ryan (1993) found that children gave more false positives in both target-present and target-absent lineups than adults. Children are as accurate as adults when the suspect is guilty but commit more false positives when the suspect is innocent (Pozzulo, Dempsey, Crescini, & Lemieux, 2009). Moreover, a study by Havard, Memon, Laybourn, and Cunningham (2012) revealed that children are better at identifying suspects of their own-age in target present procedures. Concerning the target-absent procedure, the same study showed that children made more Correct Rejections for the own-age lineup. Amongst children, the younger ones give significantly more errors than, the older children (Brigham & Van Verst, 1986).

Older adults, on the other hand, identify significantly fewer targets and more distractors than younger adults (Badham, Wade, Watts, Woods, & Maylor, 2013). Other studies also show that older witnesses produce significantly more false positives than young adults (Wise, Dauphinais, & Safer, 2007).

Emotional Status

There is a large body of evidence that points to the critical role of emotional states in memory and attention tasks (Forgas, Laham, & Vargas, 2005). Concerning to eyewitness testimony literature, emotional participants[3] provide a complete description of the perpetrator. However, they make more mistakes in recognition of the suspect (Houston, Clifford, Phillips, & Memon, 2013). A study by Valentine and Mesout (2009) showed that eyewitnesses with high state anxiety reported fewer correct descriptors, more incorrect details and made fewer accurate identifications in a lineup.

Intoxicated witness

The literature shows that when there is a target-absent lineup, witnesses that were intoxicated during the encoding (i.e., during the crime scene) are more likely to make a mistaken identification (Dysart, Lindsay, MacDonald, & Wicke, 2002). However, a study carried out by Hagsand, Roos-af-Hjelmsäter, Granhag, Fahlke, and Söderpalm-Gordh (2013) suggested that eyewitnesses who have consumed 0.4 g/kg or 0.7 g/kg ethanol have similar error rates to sober witnesses.

[3] Who viewed the emotion-inducing video.

Intellectual disabilities

Several studies have focused on the performance of people with intellectual disabilities when they have to recognize other people (Manzanero, Contreras, Recio, Alemany, & Martorell, 2012; Ternes & Yuille, 2008). Generally, the results of these studies have indicated that, although the accuracy rates are identical to those of normal individuals, false positives are significantly higher in the population with mild to moderate intellectual disability.

Race

Generally, African-American eyewitness presented better recognition accuracy than Caucasian (Brigham, Maass, Snyder, & Spaulding, 1982). The other race effect was first demonstrated by Malpass and Kravitz (1969). Effectively, accuracy rates were significantly greater for the same-race condition (Behrman & Davey, 2001). Eyewitnesses make more errors when they have to identify a perpetrator from a different race (Brigham, Bennett, Meissner, & Mitchell, 2007; Brigham & Ready, 1985). Additionally, according to Smith, Stinson, and Prosser (2004), differences were found between a same and cross-race condition in decision strategies – absolute or relative (Smith, Stinson, & Prosser, 2004). Indeed, cross-race lineups are faced using a relative judgment, whereas the same race lineups are faced using relative judgment (Smith et al., 2004). Unexpectedly, for African-American Lineups, older Caucasian adults revealed a more liberal response for sequential lineups conversely it was expected a more liberal response for simultaneous (Wylie, Bergt, Haby, Brank, & Bornstein, 2015).

The experience of contact with people from another race did not affect identification accuracy (Marcon, Meissner, Jackiw, Arbuthnott, & Pfeifer, 2008).

Sex

A study by Foster, Libkuman, Schooler, and Loftus (1994) showed that men are more influenced by lineup instructions than women. Nevertheless, as explained above, if objective and unbiased questions are used, as well as double-blind paradigms, this effect of sex ceases to exist.

Variables of the perpetrator

Stereotypes

Numerous studies have demonstrated the impact of stereotypes in categorizing people into criminals and non-criminals (Yarmey, 1993). Tellingly, some facial stereotypes have a significant impact on our ability to judge others (Yarmey, 1993).

Below we list some facial stereotypes that have been related to the identification of suspects in eyewitness testimony.

Attractiveness and distinctiveness

Eyewitness studies have focused their attention on the impact of physical attractiveness on the identification accuracy (Wells, 1978; Brigham, Maass, Snyder, & Spaulding, 1982). Wells and Olson (2003) claimed that very attractive or very unattractive faces (which are more distinctive) are easier to identify. However, low attractiveness has been linked to the criminal stereotype (MacLin & MacLin, 2004). Indeed, Saladin, Saper, and Breen (1988) found that less attractive faces are seen as more distinct, atypical and more associated with crime. Regarding the effect of distinctiveness, as early as 1974, Buckhout showed that eyewitnesses were more likely to choose an innocent from a picture array if he was distinctive from the other members.

The literature also shows that more attractive people are usually considered as more honest and, therefore, more unlikely to have committed a crime (Saladin, Saper, & Breen, 1988).

Trustworthiness and masculinity

Flowe (2012) showed that the facial features that are perceived/inferred as related to dominance and low trustworthiness could also lead to inferences of high criminality. The author also affirms that this type of judgment influences the decision-making in lineups. The results of a study conducted by Estrada-Reynolds, Reynolds, McCrea, and Freng (2016) showed that masculinity is consistently associated with criminality, particularly with violent crime.

Hairstyle changes

When the perpetrator changed his hairstyle between crime event and the lineup, the accuracy rate of the identification decreases significantly (Pozzulo & Marciniak, 2006). In fact, according to Shepherd, Davies, and Ellis (1978), the upper features of the face like hair are crucial for accurate recognition. Additionally, Pozzulo and Warren, (2003) found that hair is one of the face features by witnesses when they are describing the perpetrator and, thus, a change in it increases the likelihood of error in the identification.

Variables of the event

Violence and stress of the event

Clifford and Hollin (1981) found a positive correlation between accuracy and confidence in the identification when witnesses viewed non-violent videos. When the videos were violent, this relationship did not emerge. In the same sense, Hope, Lewinski, Dixon, Blocksidge, and Gabbert (2012) showed that arousal in a violent assault significantly reduced the accuracy in identification in the lineup. Additionally, when events are violent, peripheral information[4] is neglected when compared to neutral events (Brown, 2003; Christianson & Loftus, 1987). This absence of peripheral processing[5] makes it impossible to enhance memory by contextual reinstatement procedures because such procedures use the peripheral information to cue memory (Brown, 2003).

In the same way, as for violent events, high levels of stress during an event can cause a corrosion of memory (Deffenbacher, 1994). On the other hand, Safer, Christianson, Autry, and Österlund (1998) also showed that when the stress associated with the event increases, the attentional focus of the eyewitnesses becomes narrower (tunnel memory) which, in turn, results in a greater number of critical details recalled from a traumatic event. Christianson (1992) had already found the same result, however, in a posterior meta-analysis, Deffenbacher, Bornstein, Penrod, and McGorty, (2004) concluded that increased anxiety led to significant decreases in recall accuracy (pooled Cohen's *d* = .31).

---

[4] E.g., In Brown (2003), the automobile driver is central information and bystander is peripheral information and were identified in separate photo lineups.
[5] Depends on the theory of perception.

Light and distance

The literature shows a positive correlation between identification accuracy and illumination in the scene of the event, and a negative correlation between accuracy and the distance to the scene (Wagenaar & Van Der Schrier, 1996). A study carried out by Lindsay, Semmler, Weber, Brewer, and Lindsay (2008) showed that the accuracy of the decision declined with increasing distance. However, the distance did not affect the identification choosing rates (rate of picking a suspect independently of whether it was a target or a distractor/total of suspects). Moreover, Asai (2001) showed that 66.7% of subjects who saw the target's photo in the dark condition (absence of illumination in the scene of the event) made false positives.

Attention vs. exposure time

In this debate, the literature states that the quality of attention given to a particular crime is a better predictor of accuracy than the total time of exposure to it (Meissner, Sporer, & Schooler, 2007). The authors explain that this supremacy exists because efficient attentional processing of face characteristics increases the accuracy in identification.

Delay

Eyewitness literature claims that accuracy in lineup identifications decreases when the time between the crime and the identification in the lineup increases (Kassin, Tubb, Hosch, & Memon, 2001; Penrod, Loftus, & Winkler, 1982). For example, a study carried out by Ellis, Shepherd, and Davies (1980) revealed that participants remembered significantly fewer details after one(1) week compared with the next hour or next day condition.

After having described the variables that have been found to influence the witness, the criminal, and the event, two other relevant variables for eyewitness testimony are listed below.

Other relevant variables

Confidence in response

In 1978, Leippe, Wells, and Ostrom (1978) showed that the certainty of a choice in the lineup is not related to the accuracy of the identification. The same pattern of results was found by Hosch and Cooper (1982), who showed that eyewitnesses' accuracy was utterly unrelated to their level of confidence. According to Krafka and Penrod (1985), confidence was related to accuracy only where context was reinstated[6], and Brewer, and Wells (2006) found that the confidence-accuracy relationship is strong only for choosers.

Despite the previous results, an excellent recent study carried out by Wixted and Wells (2017), published in Psychological Science in the Public Interest, showed that there is a relationship between accuracy and confidence if we consider the fair administration of lineup.

[6] 85 Shop staff were asked to identify a customer from a photo-lineup of faces. Context was reinstated by providing physical cues from the customer.

Response latency

Dunning and Perretta (2002) have shown that when witnesses make their decisions in under 10 to 12 seconds, they will achieve about 90% accuracy. Accurate witnesses are more likely to reach their decisions automatically, without conscious thought, therefore not needing to prolong their decision time. However, Weber, Brewer, Wells, Semmler, and Keast (2004) postulate that the optimum time 10–12 s rule is not a reliable method for the identification of accurate decisions. Brewer, Caon, Todd, & Weber (2006) found a relationship between identification accuracy and response latency. However, they also establish that this evidence is not enough to ensure identification accuracy in the applied setting.

Analyzing all the previously mentioned information, we conclude that research in experimental psychology has given a set of scientific procedures to the criminal system that is necessary for obtaining reliable and accurate eyewitness evidence. However, the statement published by Levine and Tapp (1973) saying that "inaccurate identification has been and continues to be a major source of faulty convictions" (p. 1082) remains true. Indeed, errors in eyewitness remain a reality (Wells, 2014), and have become more evident with the advent of DNA analysis (Lacy & Stark, 2013). The Innocence Project[7] indicates that in 70% of the exonerations proved by the analysis of DNA, the conviction was made on the basis of the wrong identification of an eyewitness testimony (The Innocence Project, 2017). The exoneration study conducted by Gross and Shaffer (2012) analyzed in detail 873 cases in the National Registry of Exonerations and determined that 76% of these cases (667) were due to eyewitness

[7] The Innocence Project, founded in 1992 by Peter Neufeld and Barry Scheck at Cardozo School of Law, exonerates the wrongly convicted through DNA testing and reforms the criminal justice system to prevent future injustice (see https://www.innocenceproject.org/ for detailed information).

misidentifications. Due to its social impact, this type of identification errors (false positives) is the most problematic as it results in an innocent person being put behind bars (Malpass, 2006).

Facing this problem, Garry Wells (2014), a very prominent researcher in the field and broadly cited above, states that :

> "a wholly different approach probably would be developed, perhaps one involving eye movements, pupil dilation, event-related potential (ERP's) patterns, response latencies, implicit memory tests, and other potential indicia of recognition. Bringing psychological science to bear on the serious problem of eyewitness identification ought to mean much more than manipulating whether photos are shown as groups versus one at a time. The next generation of eyewitness researchers should throw out the traditional" (p. 15).

In fact, the literature is scarce in this area, particularly regarding studies that look for psychophysiological markers of recognition and accuracy.

Central Measures

On July 22, 1977, John Schweer, a retired 56-year-old police officer, appeared murdered. In November of the same year, Terry Harrington, a 17-year-old black boy, was arrested for being identified by a witness as the perpetrator and in August 1978 an all-white jury convicted Harrington to life in prison (Denzel & Possley, 2012).

Ten years later, in 1988, in a study with Event-Related Potentials (ERPs) funded by the CIA, Lawrence Farwell, supervised by Dr. Emanuel Donchin found that a P300 (a positive deflection in Electroencephalography (EEG) signal between 300 and 600ms after stimulus) is elicited with greater amplitude towards relevant stimuli to a particular subject. Facing these results, the authors postulate that P300 might be used in a forensic setting using the same method used in polygraphy, the 'guilty knowledge test'. In this procedure, the suspect visualizes a series of sequential stimuli and some of these stimuli are related to the crime and so are relevant and distinctive only to a person who knows the crime. A P300 response to these "target stimuli" works as evidence of this knowledge ("Forensic neuroscience on trial," 2001).

The defense of Terry Harrington was based in part on the P300's analysis, and Dr. Farwell did this test. In fact, the amplitude of Harrington's P300 was similar to irrelevant stimuli and the crime-related stimuli. The evidence obtained by the P300 analysis that the court called "Brain Fingerprinting Test" together with fraudulent elements in the legal process led to Harrington being released. Curiously, or not, as far as we know, the only work that studied the electroencephalographic correlates of identification accuracy using an eyewitness paradigm (Lefebvre, Marchand, Smith, & Connolly, 2007), showed that the P300 component remained a reliable predictor of accuracy in identification. In this study, four 60 second non-violent crime videos were shown to participants, who subsequently would have to identify the person responsible for each of the four crimes in the target present sequential lineup, constituted by the criminal and five suspects. The results revealed that the P300 (maximal between 400 and 600 ms after stimuli onset) was elicited with greater amplitude to the exibition of the culprit compared to the distractors. The theoretical framework behind these results is generalized from the

"Brain Fingerprinting Test" and the ‘guilty knowledge test‘, which shows that the P300 (positive potential around 300 ms after stimulus onset) is elicited with greater amplitude when we visualize a relevant stimulus (e.g., the perpetrator), compared to irrelevant stimuli (e.g., distractors). This differential elicitation only occurs for individuals with knowledge of the stimuli of the crime environment (Allen & Iacono, 1997; Farwell & Donchin, 1991; Rosenfeld, 2002). In addition to this explanation, the finding obtained by Lefevre and colleagues (2007) can be outlined within the theory of emotional processing (Lifshitz, 1966), which advocates that high arousal images (negative and positive affect) elicit a greater amplitude of the P300 than neutral images that are low arousal (e.g., Delplanque, Silvert, Hot, & Sequeira, 2005; Olofsson & Polich, 2007). In fact, the critical item (e.g., the perpetrator), being emotionally charged, elicits a larger amplitude of the P300 component (see Righi et al., 2012).

Even though there are no more studies that directly relate ERPs, neuronal patterns, or functional connectivity and localization with eyewitness testimony accuracy, there is a great body of literature that shows the psychophysiological distinction between familiar faces (e.g., the perpetrator – when he is being identified, has already been seen committing the crime) and unfamiliar faces (e.g., distractors – when they are being identified, it is the first time they are being seen), or the specific psychophysiological response to unpleasantness / arousal / negative affect faces (as the perpetrator) vs neutral faces (as distractors) (see Werner, Kühnel, & Markowitsch, 2013). This literature has shown the sensitivity of the P100 component (positive potential that peaks around 100 ms after stimulus onset) to face processing (e.g., Cunningham, Van Bavel, Arbuckle, Packer, & Waggoner, 2012; Herrmann, Ehlis, Ellgring, & Fallgatter, 2005). A study exploring the effect of facial emotion expressions on recognition memory showed an increase in the amplitude of the P100 for neutral faces that during the encoding phase were presented with a fearful expression, compared with neutral faces that originally had a neutral expression and with new faces (Righi et al., 2012). The authors point to the orbitofrontal region as a potential generator of this difference and explain that this result is due to an association between the P100 and an attentional capture of stimuli with clues of fear or threat (see Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007). Other studies with ERPs have shown that emotionally activating stimuli, negative or positive, elicit a greater amplitude of the P100 than neutral stimuli (e.g., Carretié, Hinojosa, Martín-Loeches, Mercado, & Tapia, 2004; Hot, Saito, Mandai, Kobayashi, & Sequeira, 2006). Addicionaly, Liu, Harris, and Kanwisher (2002), identified a face-selective response at approximately 100 msec that is correlated with correct face recognition. In addition to the P100 (visual sensory face processing component), also the N170 (negative deflection occurring around 170 ms after stimulus onset), which is a structural processing index of the face (see Bruce & Young (1986) model), has been shown to be sensitive to emotional facial expressions (e.g., Caharel et al., 2007). However, this result is not consensual, with some studies showing a smaller (e.g., Blau, Maurer, Tottenham, & McCandliss, 2007) or greater (e.g., Righi, 2012) amplitude of the N170 for threatening or emotionally activating faces, and other studies showing that the N170 is not affected by emotional aspects such as facial expressions (see Rossion & Jacques, 2008). Interestingly, this lack of consensus is also observed for the effect of familiarity on the N170. Indeed, some studies have shown that this component does not distinguish familiarity in faces (e.g., Eimer, 2000; Jemel, Pisani, Calabria, Crommelinck, & Bruyer, 2003), while other studies show the N170's sensitivity to familiarity (Barrett, Rugg, & Perrett, 1988; Stephanie Caharel et al., 2002).

A study carried out by Herrmann et al. (2005), with the theme "Source localization of early stages of face processing", showed that the N170 appears to be elicited in the inferior occipital cortex, in the fusiform gyrus, replicating the results found by Shibata et al., (2002). In this study, Herrmann et al. also found a hyperactivation in the medial prefrontal cortex and anterior cingulate cortex.

Regarding hemodynamic brain activity, face processing is related to fusiform regions (e.g., Kanwisher, McDermott, & Chun, 1997; Vidal et al., 2010). However, other clusters like pulvinar, inferior occipital gyrus, amygdala were found (Rossion, Hanseeuw, & Dricot, 2012). Referring to the familiar/unfamiliar face processing paradigm, Gobbini and Haxby (2006) found higher activation in the precuneus while watching familiar faces. Observing new faces led to higher responses in the fusiform gyrus and the amygdala. Von Der Heide et al. (2013) found higher left-lateralized anterior temporal lobe activations for familiar faces and a right anterior temporal activation for unfamiliar faces. Lidaka, Harada, Kawaguchi, and Sadato (2012) carried out a study attempting to investigate the neuroanatomical substrates involved in true and false memories for faces. The results show a positive correlation of familiarity with activation in the orbital cortices. False memories correlated positively with activation in the anterior cingulate cortex and amygdala. In another study on the same topic, Iidaka, Harada, and Sadato, (2014), found a cluster in the medial part of the prefrontal cortex associated with false face memories. Similarly, Van Kesteren and Brown (2014) suggested that, despite the importance of the medial prefrontal cortex for memory and decision making, these areas are related to false memories or false confidence (when we are convinced, but we are wrong) about some details of memories. Concerning the emotion processing paradigm, faces with emotional cues led to a hyperactivation of the amygdala when compared to neutral faces (Adams, Gordon, Baird, Ambady, & Kleck, 2003; Adolphs, 2008). An amygdala–hippocampal cluster activation was found in the perception of fearful faces (Phillips et al., 2004). A study by Reinders et al., (2006) established that the latency of the BOLD (blood-oxygen-level dependent) activity in the amygdala-hippocampal cluster (bilaterally) was found to be more than 500 ms earlier for fearful faces than for neutral faces.

Peripheral Measures

Peripheral measures of the central nervous system have been used in the forensic context since the beginning of the last century. In 1914, Vittorio Benussi, the Italian psychologist found a relationship between lying and breathing, and the following year the American lawyer and psychiatrist William Marston found a relationship between heart pressure and lying (Synnott, Dietzel, & Ioannou, 2015). However, in spite of these initial advances, the first device that allowed simultaneous synchronization of blood pressure heart rate and respiration rate, which was called "the polygraph", was created in 1921 by physiologist John A. Larson (Hyman, 1989). The debut of the polygraph in court was not the best, and US Supreme Court banned its use for lacking scientific evidence ("Forensic Neuroscience on trial," 2001). Later, history shows an "in and out" of the polygraph in the American judicial system, such as Ronald Reagan's 84th directive, which authorized all federal agencies to use polygraphs, which was revoked less than three months later (Synnott et al., 2015). Despite this, the use of polygraph with the 'guilty knowledge test' paradigm presents good scientific evidence (Gamer, Verschuere, Crombez, & Vossel, 2008; Klein Selle, Verschuere, Kindt, Meijer, & Ben-Shakhar, 2016; Verschuere, Crombez, Smolders, & De Clercq, 2009).

Although the aforementioned peripheral measures are only of autonomic etiology, the ocular movement patterns, which have a somatic origin, are also used in a criminal context, for example in the signature analysis (Dyer, Found, & Rogers, 2006). Specifically, in the eyewitness field, there are only a few works that use eyetracking[8] to study eye movement patterns associated with accurate identifications and decision-making processes in lineups (e.g., Flowe, 2011; Flowe & Cottrell, 2011; Mansour, Lindsay, Brewer, & Munhall, 2009). These studies show that "face dwell time" (total time looking at the face) varies according to the lineup members (e.g., Flowe & Cottrell, 2011). Flowe (2011), in a study comparing sequential and simultaneous lineups, showed that participants facing a sequential lineup spend more time looking at each of the lineup elements. According to this author, participants who saw sequential lineups appeared to perform a more extensive examination of the faces than participants in the simultaneous lineup because they only had one opportunity to see each particular face. The author concludes that sequential lineups require a greater degree of correspondence between a face and the image of the perpetrator of the crime that the eyewitness has in memory before making a positive identification. In the same study, it has also been observed that participants who saw the sequential lineup spent more time looking at the external variables of the face than those in the simultaneous lineup condition. This result assumes greater importance for the fundamental role that the

[8] Technique whereby a participants' eye movements are measured, where we can have, for example, the specific place where the participant is looking (Poole & Ball, 2005). The point-of-regard can be found through the trigonometric analysis of corneal reflection to infrared light emitted by the device (Duchowski, 2003).

external characteristics of the face have for the recognition of unfamiliar faces (Flowe, 2011; Young, Hay, McWeeny, Flude, & Ellis, 1985). Indeed, considering the unfamiliar faces, external features, when shown alone, are more important to recognizing faces than when only internal features are shown (Ellis, Shepherd, & Davies, 1979).

Additionally, Flowe and Cottrell (2011) showed that for the simultaneous lineup, when a distractor was incorrectly identified, participants made a greater number of visits to the distractor than to the target when the target was identified and, thus, concluded that the deliberative process occurs with more intensity when the signal/mnesic trait is low. About the number of visits, Flowe and Cotrell's study postulated that when the target was absent, the number of visits to the unidentified distractors was significantly higher when the lineup was rejected without identification (correct rejection), than when a distractor was identified (false positive). The fact that there is not enough evidence lead all these researchers to recommend caution in generalizing the results.

Regarding autonomic measures, as far as we are aware, there are no studies that relate these measurements to eyewitness testimony. According to Sokolov (1963), from an autonomic perspective, the significance and novelty of a stimulus models a process called oriented reflex. This orienting response has received much attention regarding the study of its autonomic correlates. Indeed, emotion and stimuli significance (positive or negative affect) play an important role in cardiac response (e.g., Bradley, 2009; Lang, Bradley, & Cuthbert, 1997) and in facial muscle activity (e.g., Dimberg, 1982; Tan et al., 2016). Regarding heart rate, Campbell et al. (1997) coined the term "fear bradycardia", which is a typical pattern of directed reflex in all mammals, manifested by a deceleration in heart rate when the organism is confronted with threatening cues. In this sense, the literature has shown that cardiac deceleration is significantly higher for unpleasant images compared to neutral or even pleasant ones (e.g., Bradley, 2009; Bradley & Lang, 2000). Lie-detection studies also showed a heart rate deceleration towards significant stimuli (Gamer et al., 2008; Selle et al., 2016). Also, heart rate acceleration occurred when people are facing happy faces and deceleration when they are facing angry faces (e.g., Johnsen, Thayer, & Hugdahl, 1995; Palomba, Angrilli, & Mini, 1997).

Regarding electromyography, this technique has been considered a robust method to recognize affective states or different emotions (Cacioppo, Petty, Losch, & Kim, 1986; Tan et al., 2012; Tan et al., 2016). Schwartz and his colleagues found that, whereas unpleasant imagery elicited greater activity over the corrugator supercilii, pleasant imagery elicited greater activity over the zygomaticus major (Brown & Schwartz, 1980; Schwartz, Fair, Salt, Mandel, & Klerman, 1976). The same result was found for pleasant and unpleasant faces (Dimberg, 1990). Tellingly, the activity over the corrugator supercilii is inversely related to the valence of a subjective experience (Kawamoto, Nittono, & Ura, 2013; Larsen, Norris, & Cacioppo, 2003), is associated with negative affect (Cacioppo, Martzke, Petty, & Tassinary, 1988; Dimberg, 1982) and represents a motivation to withdraw (see Davidson, 1995; Davidson, Ekman, Saron, Senulis, & Friesen, 1990) from a stimulus (Allen, Harmon-Jones, & Cavender, 2001). In turn, zygomatic activity is directly associated with positive valence (Ribeiro, Teixeira-Silva, Pompéia, & Bueno, 2007) and related to positive affect (Cacioppo et al., 1988), which is associated with an approach motivation (Davidson, 1990).

Given the aforementioned, and due to the scarcity of evidence that supports the existence of psychophysiological markers of performance in eyewitness testimony, this doctoral dissertation focused on providing evidence about this gap.

Specifically, this work aimed to identify neuropsychophysiological patterns of identification accuracy in eyewitness testimony and, using eyetracking, try to bring new perspectives into the sequential vs. simultaneous lineup debate. In order to achieve this, we developed one pilot study and three empirical studies which are described in the structure below.

In terms of structure, the present thesis is organized into six chapters. Chapter 1 presents a review of the main literature relevant to the research area of eyewitness testimony, including the new methodologies (central e peripheral psychophysiological measures and occulometry) adopted in the present work. Chapter 2 describes the pilot study that allowed us to construct the materials for the three empirical studies. Chapter 3 reports a study investigating the patterns of ocular movements in eyewitness testimony. Chapter 4 reports a study integrating central and peripheral measures, aiming to identify psychophysiological markers of eyewitness performance regarding electroencephalography, facial electromyography, and electrocardiography. Chapter 5 presents a study exclusively with peripheral measures, where the electromyographic and electrocardiographic indicators that emerged in the previous study are explored in more detail. Finally, Chapter 6 presents a general discussion of all the work, integrating the results from the three empirical studies with the literature in eyewitness testimony, discusses the contribution of this work to this important field and provides some suggestions for directions for future work.

CHAPTER II

DEVELOPMENT OF THE EXPERIMENTAL MATERIALS: FILMING OF THE CRIME VIDEOS AND SELECTION OF THE PHOTOGRAPHS FOR THE LINEUPS

DEVELOPMENT OF THE EXPERIMENTAL MATERIALS

600 male faces from NOTTINGHAM[9], PIE[10], TEXAS[11], UBI[12] and FEI BRAZILIAN[13] databases were selected. All faces were in frontal position and displayed a neutral expression, and all images were standardized for color (grayscale), luminosity, grain and size (360 * 260px) using a Paint Shop Pro X4 script.

Further development of the experimental materials was carried out in 3 phases, which are described in detail below.

1st Phase – Filming

16 crime situations were recorded in video: eight of theft (goods were stolen without confrontation with the victim) and eight of robbery (there was confrontation with the victim); In the same type of scenario (when the victim was withdrawing money at an ATM cash point) and with a camera tripod in the same position and angle to the ATM cash point. Each of the eight hired actors - all selected from a pool of actors of the GRETUA - Experimental Theatre Group of the University of Aveiro - performed one theft situation and one robbery situation. In the videos, the actor's face remained in frontal position for 4 seconds. On the same day of filming, the actors took a photograph, also in frontal position, to include in the face database. Subsequently, all eight actors were evaluated in terms of attractiveness and distinctiveness by ten men and ten women, using a 1 to 7 Likert scale. All scored between 3 and 4.99.

---

[9] Retrieved from http://pics.psych.stir.ac.uk/2D_face_sets.htm
[10] Database no longer available, was replace by http://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Home.html
[11] Database no longer available, was replace by http://live.ece.utexas.edu/research/texas3dfr/
[12] Private Database - Photos captured within the doctoral thesis “O efeito do dimorfismo sexual e da confiabilidade percebida nas preferências de atratividade: uma perspetiva evolutiva” by Mariana Carrito.
[13] Retrieved from http://fei.edu.br/~cet/facedatabase.html

2nd Phase – Attractiveness and Distinctiveness Evaluation

30 participants (15 women) aged between 18 and 30 years, were confronted with photographs of the faces of 600 individuals from the database and rated each face on its attractiveness on a Likert scale (1-7 - See Figure 1, left panel). Another sample of 30 participants (15 women) completed a similar task, but judged the distinctiveness of the faces (1-7 - See Figure 1, right panel). For both procedures (evaluation of attractiveness and distinctiveness) the OpenSesame software (Mathôt, Schreij, & Theeuwes, 2012; see http://osdoc.cogsci.nl) was used. Stimuli (the 600 faces) were shown sequentially and in random order, and the response was given by pressing the numeric keypad number that best corresponded to its evaluation (between a minimum of 1 and a maximum of 7).

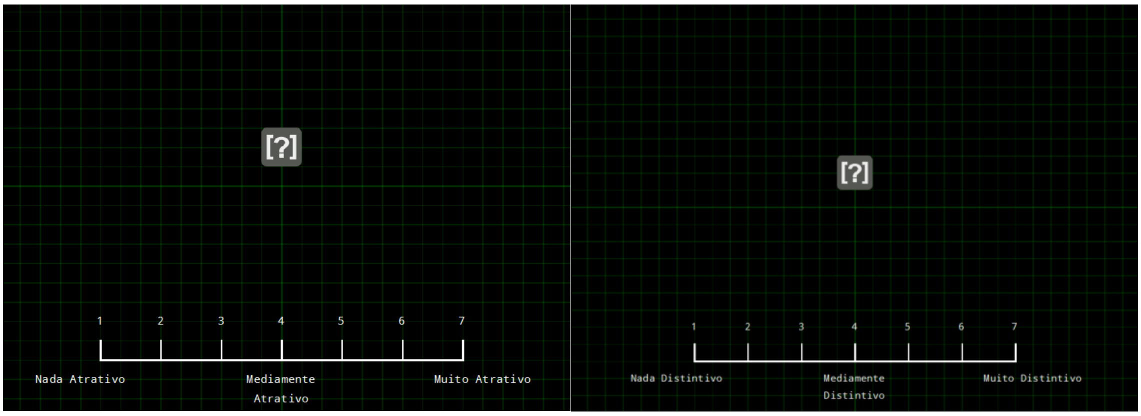


Figure 1. Print screens of the attractiveness and distinctiveness evaluation tasks in OpenSesame (Left Panel: 1=not attractive; 4= averagely attractive; 7=very attractive; Right Panel: 1=not distinctive; 4= averagely distinctive; 7=very distinctive)

3rd Phase - Similarity Evaluation

In the third phase, 40 participants (20 women) aged between 18 and 30 years, performed a similarity judgment task between each face of the database (600 faces) and the actors' (8 faces). In this task (see Figure 2), participants had to respond, on a Likert scale from 1 to 7, how similar the faces appearing on the right side of the screen were to the face on the left (belonging to the actor).

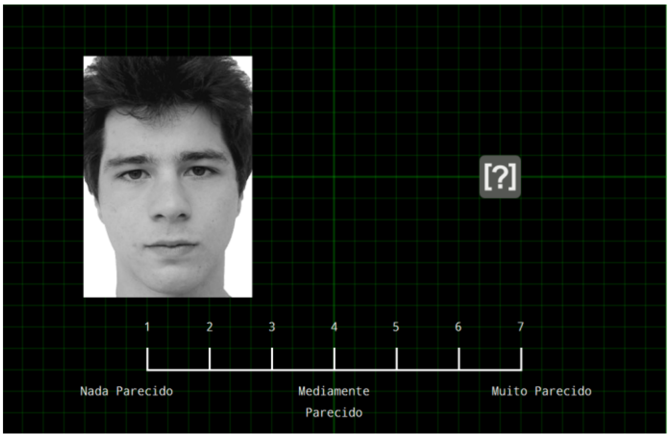


*Figure 2.*Print screen of the similarity evaluation task (1=not similar; 4= averagely similar; 7=very similar)

From the 3 Phases described previously resulted 16 videos (eight of theft and eight of robbery) and 16 lineups (eight with target present and eight with target absent). Each lineup consisted of six photographs: one actor + five distractors (in the target present condition); or six distractors (in the target absent condition). By using a photovoltaic cell, the selected image pool was measured for luminosity. Figures 3, 4, 5, and 6 illustrate the sets of face photographs selected for each crime, where the values of luminosity, attractiveness, and distinctiveness are shown for both the target and the six distractors (an average of the 6). Average values of similarity to target are also shown for the group of distractors[14]. By splitting the 7-point Likert scale into three equal parts, we defined the cut-off points to classify the faces regarding attractiveness, distinctiveness, and similarity:

1 to 2,99 – Low attractiveness/distinctiveness/similarity;

3 to 4,99 – Medium attractiveness/distinctiveness/similarity;

5 to 7 – High attractiveness/distinctiveness/similarity.

Using the related samples Sign Test, none of the variables registered significant differences between targets and distractors: luminosity, Z = -.364, p = .727[15] (see Figure 11, left panel), attractiveness, Z = -.1.061, p = .289[16] (see Figure 11, central panel), and distinctiveness, Z = -.364, p < .727[17] (see Figure 11, right panel). Due to the unlikely equality of the Z statistic and the probability value for luminosity and distinctiveness, the Wilcoxon Signed Ranks Test values were also calculated (Luminosity – Z = -700, p = .484; Attractiveness – Z = -1.122, p <. 262; Distinctiveness – Z = -.701, p < .483) and the same pattern of results was found.

[14] When the target is present, the withdrawn distractor is one that has obtained values that are farther from the target.
[15] Despite the absence of statistically significant differences between targets and distractors, the reduced number of participants increases the type II error, and therefore these results must be read carefully.
[16] Idem footnote #15.
[17] Idem footnote #15.

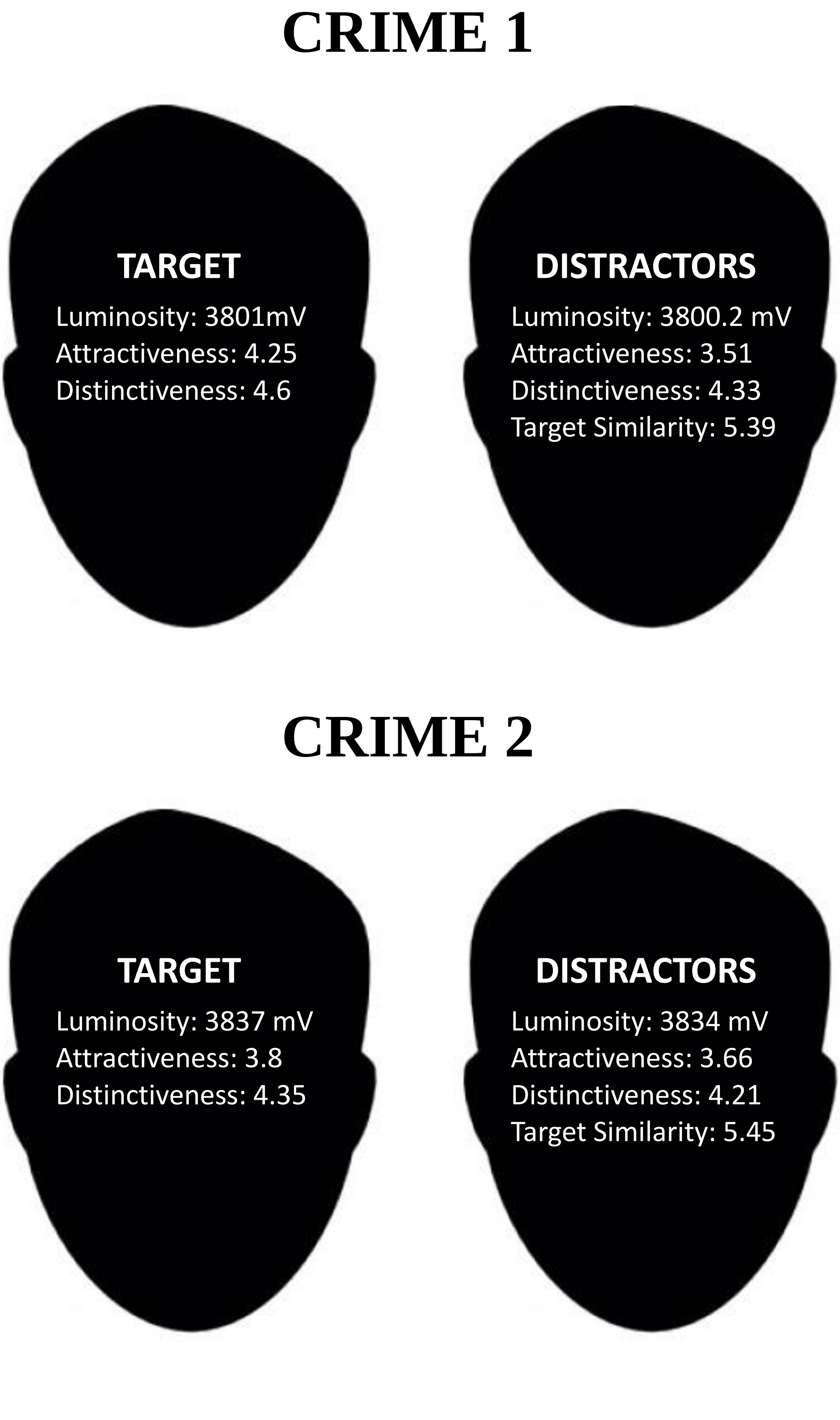


Figure 3. Luminosity, attractiveness, distinctiveness and target similarity (distractors) for individuals included in CRIME 1 and 2 lineups.

Figure 4. Luminosity, attractiveness, distinctiveness and target similarity (distractors) for individuals included in CRIME 3 and 4 lineups.

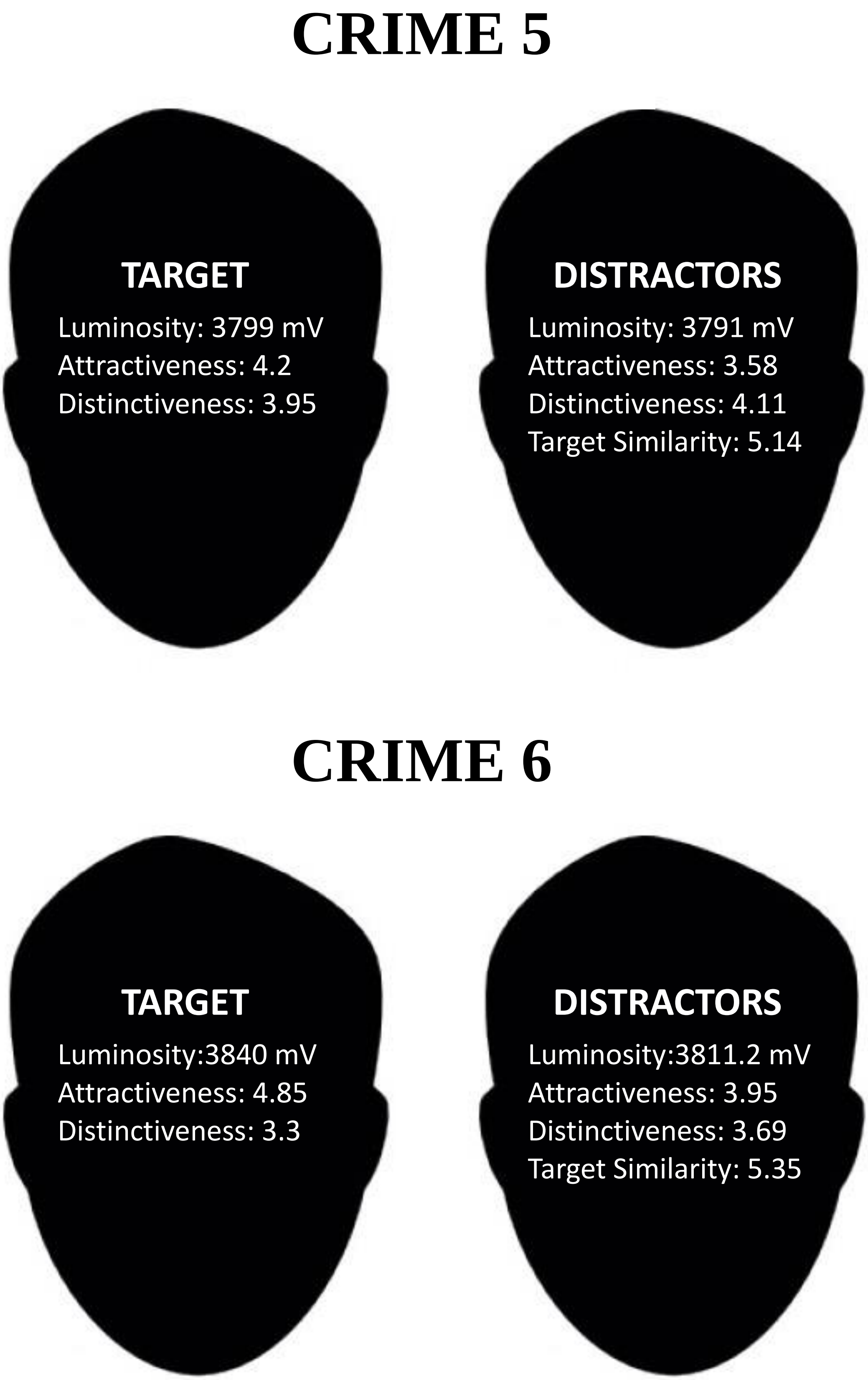
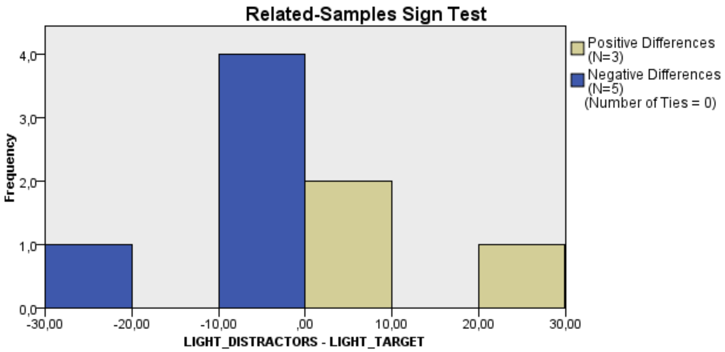


Figure 5. Luminosity, attractiveness, distinctiveness and target similarity (distractors) for individuals included in CRIME 5 and 6 lineups.

Figure 6. Luminosity, attractiveness, distinctiveness and target similarity (distractors) for individuals included in CRIME 7 and 8 lineups.

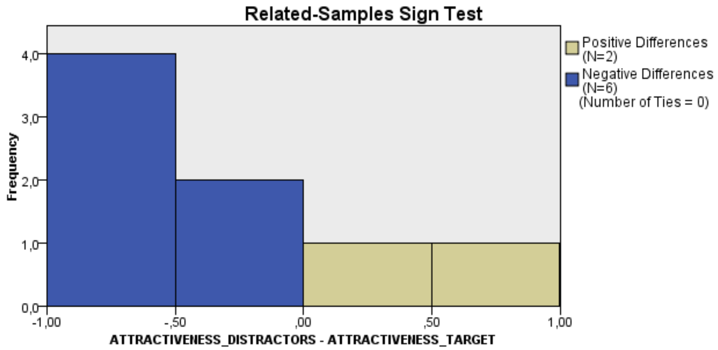
LUMINOSITY DIFFERENCES



Total N	8
Test Statistic	3,000
Standard Error	1,414
Standardized Test Statistic	-,354
Asymptotic Sig. (2-sided test)	,724
Exact Sig. (2-sided test)	,727

1. The exact p-value is computed based on the binomial distribution because there are 25 or fewer cases.

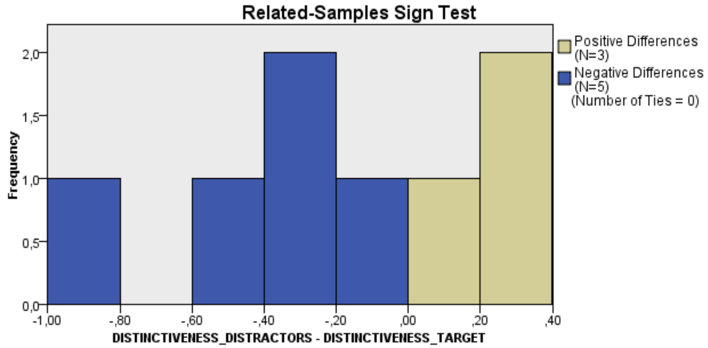
ATTRACTIVENESS DIFFERENCES



Total N	8
Test Statistic	2,000
Standard Error	1,414
Standardized Test Statistic	-1,061
Asymptotic Sig. (2-sided test)	,289
Exact Sig. (2-sided test)	,289

1. The exact p-value is computed based on the binomial distribution because there are 25 or fewer cases.

DISTINCTIVENESS DIFFERENCES



Total N	8
Test Statistic	3,000
Standard Error	1,414
Standardized Test Statistic	-,354
Asymptotic Sig. (2-sided test)	,724
Exact Sig. (2-sided test)	,727

1. The exact p-value is computed based on the binomial distribution because there are 25 or fewer cases.

Figure 7. Differences between targets and distractors for Luminosity, Attractiveness, and Distinctiveness using the Related Samples Sign Test.

Saliency analysis

Amongst other studies, the stimuli presently created would be used on an eyetracking study (see Chapter III), where it would be necessary to analyse variables such as time to the first fixation, which are known to be influenced by stimulus saliency (Itti & Koch, 2000; Itti, Koch, & Niebur, 1998). Therefore, we also intended to check whether the target was more salient than the distractors in the target present simultaneous lineup and whether the total saliency of the simultaneous lineup with target absent was less than the total saliency of the target present simultaneous lineup. For this, saliency maps were created for the 16 simultaneous lineups (eight Target Present and eight Target Absent).

Using the Saliency Toolbox for Matlab, created by Walther and Koch (2006), the saliency maps were calculated based on the Intensity (Weight = 1), Orientation (Weight = 1) and Skin hue (Weight = 1), with local max normalization. The two saliency maps for the CRIME 1 lineups (Target Present and Target Absent), as well as the two winner-takes-all coordinates[18], are shown in Figure 8.

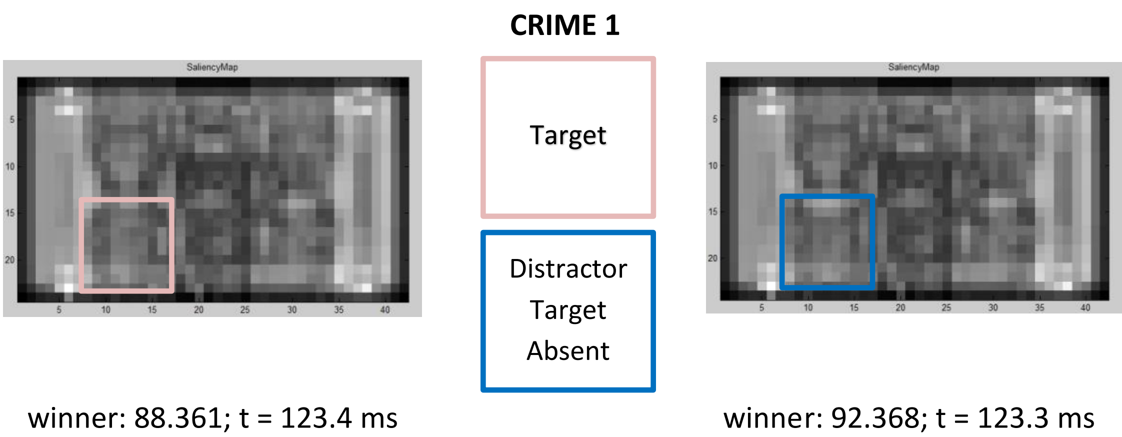


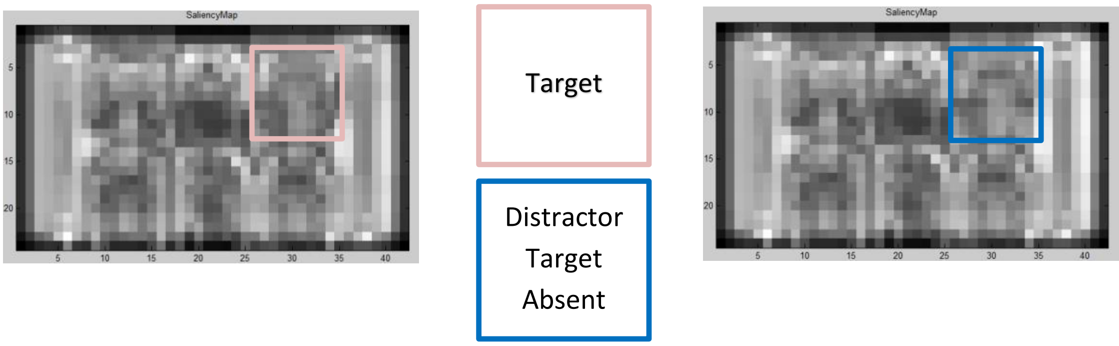
Figure 8. Saliency maps and Winner-takes-all coordinates for target present (Left panel) and Target absent (Right panel) lineups in CRIME 1

[18] Winner-takes-all (WTA) detects the most salient location and directs attention towards it. The displayed values are the coordinates of the most attended location in the image(winner) and the simulated time(t) that it took to attend to this location (Itti & Koch, 2000; Walther & Koch, 2006).

Comparing the two panels, we can see that the values of the winner-takes-all coordinates are very similar. Additionally, there are neither large saliency differences between the target and the distractors in the images nor between the total saliency of the target present lineup and target absent lineup[19].

The two saliency maps for the CRIME 2 lineups (Target Present and Target Absent), as well as the two winner-takes-all coordinates, are shown in Figure 9.

CRIME 2



winner: 94.362; t = 168.9 ms

winner: 93.365; t = 169.2 ms

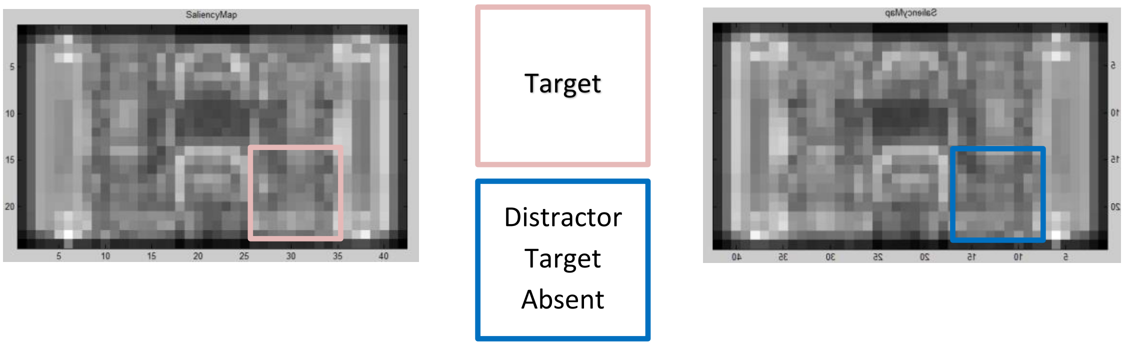
Figure 9. Saliency maps and Winner-takes-all coordinates for target present (Left panel) and Target absent (Right panel) lineups in CRIME 2.

The values of the winner-takes-all coordinates are very similar. In fact, although there is not much difference between the target and the distractors, the target seems to be the least salient. No big differences are noticeable between the total saliency of the target present lineup and the absent target lineup.

[19] Lineup images are merely illustrative, since the position of all Lineup actors (targets and distractors) was randomized.

The two saliency maps for the CRIME 3 lineups (Target Present and Target Absent), as well as the two winner-takes-all coordinates, are shown in Figure 10.

CRIME 3



winner: 88.358; t = 146.3 ms

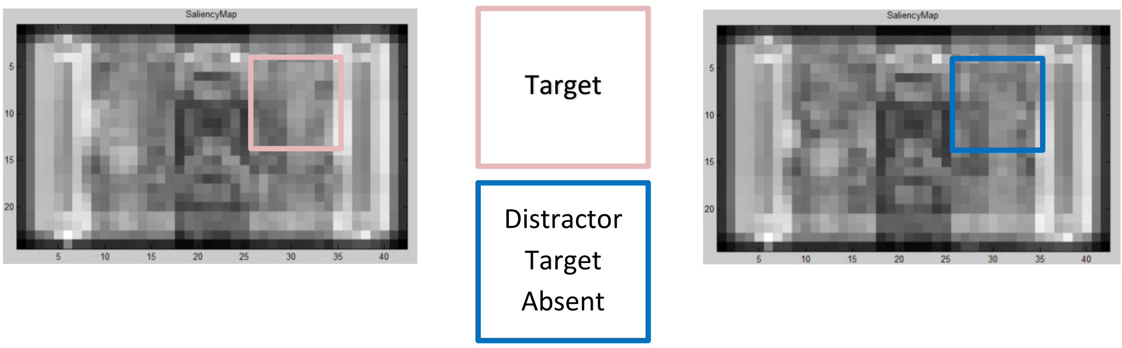
winner: 91.363; t = 147.5 ms

Figure 10. Saliency maps and Winner-takes-all coordinates for target present (Left panel) and Target absent (Right panel) lineups in CRIME 3.

Again, the values of the winner-takes-all coordinates are very similar, and no major differences between the total saliency of the target present lineup and target absent lineup were found. Despite there being no major differences between the target and other distractors, there are two distractors that stand out.

The two saliency maps for the CRIME 4 lineups (Target Present and Target Absent), as well as the two winner-takes-all coordinates, are shown in Figure 11.

CRIME 4



winner: 82.363; t = 146.6 ms

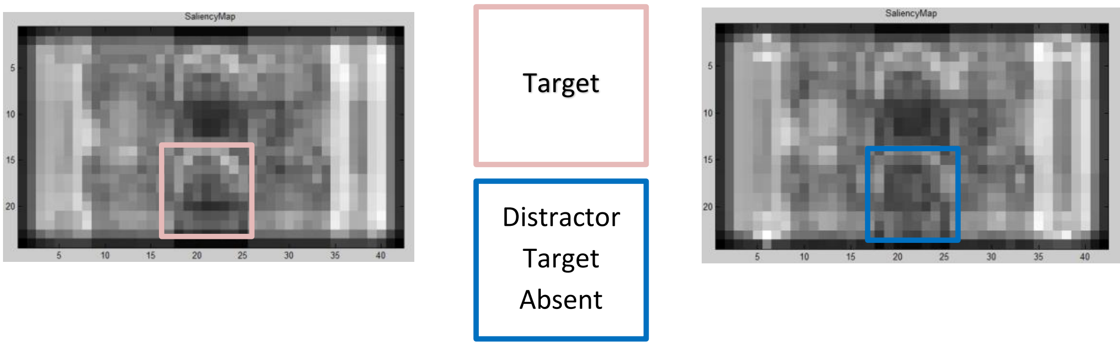
winner: 84.356; t = 146.4 ms

Figure 11. Saliency maps and Winner-takes-all coordinates for target present (Left panel) and Target absent (Right panel) lineups in CRIME 4.

The values of the winner-takes-all coordinates are once again very close. No large saliency differences were visible between the target and the distractors, in the images. No big differences between the total saliency of the target present lineup and target absent lineup were found.

The two saliency maps for the CRIME 5 lineups (Target Present and Target Absent), as well as the two winner-takes-all coordinates, are shown in Figure 12.

CRIME 5



winner:91,45; t = 143.6 ms

winner: 88,364; t = 142.8 ms

Figure 12. Saliency maps and Winner-takes-all coordinates for target present (Left panel) and Target absent (Right panel) lineups in CRIME 5.

Similarity in winner-takes-all coordinates values was once again achieved. No visual differences between the target and the distractors were found. The same happened between the total saliency of the target present lineup and target absent lineup.

The two saliency maps for the CRIME 6 lineups (Target Present and Target Absent), as well as the two winner-takes-all coordinates, are shown in Figure 13.

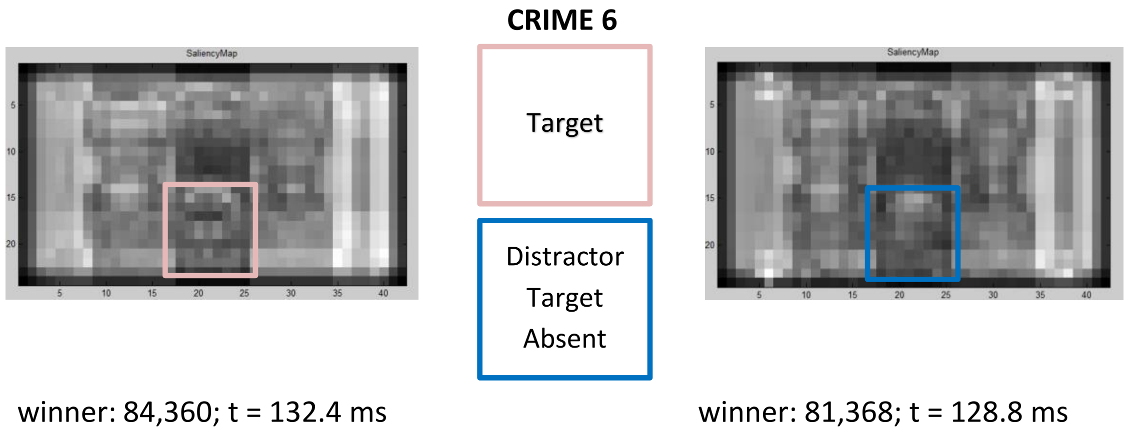


Figure 13. Saliency maps and Winner-takes-all coordinates for target present (Left panel) and Target absent (Right panel) lineups in CRIME 6.

Likewise, the values of the winner-takes-coordinates are very close, and no significant differences between the total saliency of the target present lineup and target absent lineup were registered. The saliency of target and distractors were also similar.

The two saliency maps for the CRIME 7 lineups (Target Present and Target Absent), as well as the two winner-takes-all coordinates, are shown in Figure 14.

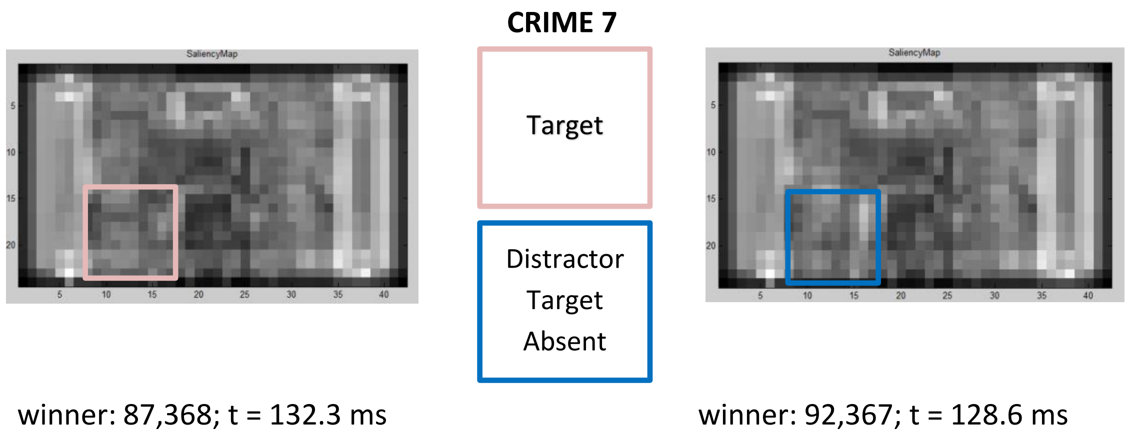
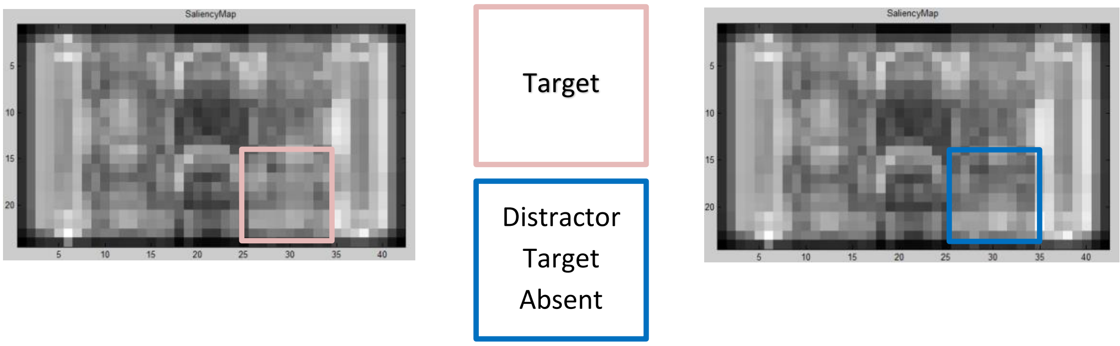


Figure 14. Saliency maps and Winner-takes-all coordinates for target present (Left panel) and Target absent (Right panel) lineups in CRIME 7.

Very similar values of the winner-takes-all coordinates were obtained. Importantly, although not much difference between the target and the distractors was observed, there was one distractor with higher saliency level. No major differences were found between the total saliency of the target present lineup and target absent lineup.

The two saliency maps for the CRIME 8 lineups (Target Present and Target Absent), as well as the two winner-takes-all coordinates, are shown in Figure 15.

CRIME 8



winner: 86,356; t = 147.9 ms   winner: 83,369; t = 147.5 ms

Figure 15. Saliency maps and Winner-takes-all coordinates for target present (Left panel) and Target absent (Right panel) lineups in CRIME 8.

Considering the winner-takes-all coordinates values present in each panel of Figure 15, we observe that results were very close. There are no visible differences in saliency between the target and the distractors in the images and between the total saliency of the target present lineup and target absent lineup. Yet, the central distractors (2 distractors in the center of the lineup) showed a higher saliency level.

Concluding Remarks

The experimental materials created have been shown to be reliable for use in the following psychophysiological studies, and seem to ensure that there are no differences in attractiveness, distinctiveness, luminosity, and low-level features between the several stimuli (targets and distractors) and between target present and target absent lineups.

Yet, as the analysis of salience measures was based on low-level features (Walther & Koch, 2006), and assuming a strict and conservative stance, the small differences between targets and distractors in some crimes can represent problems of inference in the analysis of a fast attention-grabbing variable that is the time until first fixation (Naber & Nakayama, 2013). Faced with this fact, in chapter III, where the methodology of eyetracking is used, this ocular metric (time until the first fixation) is not displayed.

CHAPTER III

GAZE PATTERNS IN EYEWITNESS TESTIMONY

INTRODUCTION

Eyewitnesses are often decisive in solving crimes and are, sometimes, the only available source of information to help determine the identity of the perpetrator (Wells & Olson, 2003). However, a large body of evidence has demonstrated the lack of reliability of eyewitness testimony (e.g., Busey & Loftus, 2007; Kassin & Gudjonsson, 2004). In fact, about 70% of people who were exonerated by DNA analysis were involved in cases of incorrect eyewitness identifications (Innocence Project, 2017). In order to try to reduce recognition errors, several procedures are used, including lineups (Busey & Loftus, 2006). For that reason, since the mid-1970s, eyewitness evidence obtained through lineups has been the focus of numerous reliability tests (Wells, 2014). The two most common lineup types are the simultaneous lineup (SIML) and the sequential lineup (SEQL) (McQuiston-Surrett, Malpass, & Tredoux, 2006). In the SIML, the members of the lineup (usually six people, being one suspect and five fillers) are presented at the same time (Wogalter, Malpass, & McQuiston, 2004). According to the Police Executive Research Forum (2013), this lineup type is the most commonly used in the United States. However, critics of this procedure claim that it encourages relative judgment since it forces the subject/eyewitness to compare all members of the lineup, resulting in the identification of the suspect that is most similar to the culprit, even when the culprit is not present (Wells, 1993). Lindsay and Wells (1985) have developed the SEQL procedure – the most commonly used type of lineup in Canada (Beaudry & Lindsay, 2006). In this procedure, suspects are presented sequentially, which reduces the problem of relative judgment by promoting an absolute judgment. Here, the witness compares each element of the lineup with his or her memory of the perpetrator and not with the other members of the lineup. Amongst the important instructions given in SEQL, the "stopping rule" is the key element (although almost never done in practice) of this type of procedure (Linsay & Wells, 1985). This rule, as the name suggests, implies that, as soon as a positive identification of the suspect of the crime occurs, images stop being shown, preventing relative judgment and promoting certainty in identification. Also, in a meta-analysis (Steblay et al., 2011) weighing 72 studies (13143 participants) in which the two types of lineup were compared, the supremacy in reducing errors of SEQL was demonstrated. This was largely explained by the fact that the use of absolute judgment decreases the number of **False Alarms**/Positives (Steblay et al., 2001; Pozzulo et al., 2015), for it promotes a conservative choice (Meissner et al., 2005). However, if it is true that a SEQL provides fewer **False Alarms**/Positives, it also provides a lower rate of culprit identification (e.g., Mecklenburg, 2006; Mansour & Flowe, 2010; Levi, 2016). Indeed, using the ROC analysis, a method recommended by Gronlund, Wixted and Mickes (2014) to verify which type of lineup is superior, the results of several studies show a superiority of the SIML, being diagnostically superior in identifying suspects (e.g., Amendola & Wixted, 2015; Carlson & Carlson, 2014; Mickes, 2015; Wixted & Mickes, 2014).

Faced with this ambivalence, Malpass (2006) introduced some apparently inconclusive moral and political notions into the debate: should we increase the probability of identifying the perpetrator at the risk of increasing the possibility of mistaken identifications? Or should mistaken identifications be avoided at all costs, enabling a criminal to escape sentence? In the same sense, Wells (2014) argues that the question that is implied is "Should we adopt a new procedure (simultaneous) that increases the chances that the guilty might be identified, but also increases the chances of mistaken identification?" (p. 14).

None of the lineup types assumes a clear superiority, so Demarchi (2013) argues that we must use both methods as complementary and that, taking into account the characteristics of the investigation, we must choose the appropriate type of lineup. If the suspect is arrested shortly after the crime and there is very little (if any) incriminating evidence (which implies a significant margin of error), SEQL is preferable. On the other hand, if the prosecution has a significant amount of incriminating evidence, SIML is advisable. In the same way, a crime with a larger criminal frame (number of years in reclusion) should be advised the use of a SEQL.

In 2014, Garry Wells, the most influential name in eyewitness literature, stated that, in the face of this impasse (SEQL vs. SIML), different approaches should be used to improve identification accuracies, such as eye movements, pupil dilation, event-related potential patterns, response latencies and implicit memory tests. Regarding eye movements, which is the focus of this chapter, since the 1980s, that the tracking of eye movements has been applied to the study of memory in forensic contexts (e.g., Loftus, Loftus, & Messo, 1987). However, only recently researchers began to study decision-making processes in lineups using eyetracking (e.g., Flowe, 2011; Flowe & Cottrell, 2010; Mansour, Lindsay, Brewer, & Munhall, 2009). The first attempt to study the visual behavior in lineup performance was made by Mansour, Lindsay, Brewer, & Munhall, in 2009. This study that used SIML showed that face dwell time (total duration of time that participants fixate a face, during a lineup) was longer for the positively identified face, independently of whether it was the target (correct response) or a distractor (false positive/alarm). It was also shown that the participants took longer to make a decision when they rejected the lineup and that when participants identified a target, they spent less time on visual exploration than when they identified a distractor. Additionally, the authors found that fewer comparisons were made when the target was correctly identified and that numerous comparisons may indicate that the lineup does not contain the criminal. The decision time results showed that a long time to make a decision was related to the **False Alarms**/Positives, consistent with previous literature (Dunning & Stern, 1994). Although these results were promising, the effect sizes were greatly reduced, leading the authors to conclude that visual behavior is a weak predictor of identification accuracy.

In a study comparing SEQL and SIML, Flowe (2010) showed that participants facing a SEQL had a longer dwell time at each of the lineup elements than participants in the SIML. Additionally, the faces that were not positively identified were evaluated for longer in the SEQL, but the faces that were positively identified registered similar times between lineup types. According to these authors, it appears that participants who saw a SEQL performed a more extensive examination of the faces than participants in the SIML, maybe because they only had one opportunity to see a particular face. The author concluded that to make a positive identification, SEQLs require a greater degree of correspondence between the memory of the eyewitness and the actual face of the perpetrator. In the same study, it was demonstrated that participants who saw the SEQL spend more time looking at the external features of the face than those in the SIML condition. This result shows that the SEQL seems to be more sensitive and denote better the fundamental role that the external features of the face have in recognition of unfamiliar faces (Young, 1985).

The results obtained by Flowe and Cottrell (2011) showed once again that participants in the SIML dwell longer on the identified faces, confirming Mansour et al.'s (2009) findings. It was also shown that, for the SIML, when a distractor is incorrectly identified, participants give/perform a greater number of visits to that distractor than to the target, when he/she correctly identified. This suggests that a greater deliberative process occurs when the mnesic signal/trace is low. Flowe and Cottrell (2011) verified that when the target was absent, the number of visits to the unidentified distractors was significantly higher when the lineup was rejected (correct rejection) than when a distractor was identified (false positive).

Previous studies showed that when witnesses see a second lap of a SEQL, the number of **False Alarms**/Positives significantly increases (Lindsay, Lea, & Fulford, 1991; Steblay et al., 2011). Therefore, this does not seem to be a good way to increase identification accuracy. This study attempts to test a different methodology, combining the two types of the lineup with the objective of improving performance and uses eyetracking data to investigate differences between eye movement patterns in the various conditions and how they relate to identification accuracy. Thus, this study aimed to verify if a procedure consisting of a SEQL followed by an SIML allows taking advantage of the best features/potential of the two lineups. That is, we aimed to understand if the SEQL in a first phase allows the reduction of the number of **False Alarms**/Positives and an SIML in the second phase allow to maintain the judgment of the first phase, and increase the number of correctness in identifying the culprits. Regarding the eyetracker variables, we want to explore this advantage through a carryover effect (Transfer of characteristics from the first to the second phase) using the number of fixations and dwell time. Using this eyetracking measures we also want to try to respond to Wells' (2014) challenge and explore potential markers of correct identifications, taking into account the type of lineup and the presence of the target. Due to the lack of literature and the aforementioned absence of robust findings, we also tried to replicate some of the results found in previous eyetracking studies.

METHOD

Participants

Sixty undergraduates (30 women), with mean age of 21.13 years (*SD* = 2.22), were divided randomly (conditional) into three groups. All groups consisted of 20 participants (10 females), and the age range in Group 1 was between 18 and 25 years old (*M* = 20.20; *SD* = 1.79), between 18 and 27 (*M* = 21.20; *SD* = 2.40) in Group 2, and between 18 and 28 (*M* = 22.00; *SD* = 1.5) in Group 3. The participants who wore lenses, glasses and had visual problems or pathologies were excluded. Eleven participants were also excluded due to calibration problems.

Materials

Videos and photos

Eight theft videos recorded at an ATM were used. The recording of these vídeos was described in detail in the previous chapter. As previously mentioned, in all these videos the face of the culprit was presented in frontal view for 4 of the 20 seconds of the total duration of the video. Also, all face images used in the lineups described below were in grayscale (360 * 260px), rated with high similarity to the culprit of the video, and rated with average attractiveness and distinctiveness.

Lineup procedures

Three types of lineup procedures were designed especially for this study: A, B, and C. Procedure A was completed by group 1, procedure B by group 2 and procedure C by group 3.

PROCEDURE A consisted of a classic SIML with 6 faces, either in a target present (five distractors and the culprit seen in the video), or in a target absent (6 distractors) condition, arranged simultaneously in a frame, with three faces on top and three faces on the bottom, presented in a randomized order. Above each photo, there was a number from 1 to 6. Participants had to check if any of the faces presented in the lineup was the person who committed the crime. To do so, using the numeric keypad, participants would have to press the key with the number above the selected picture (1, 2, 3, 4, 5, or 6), or press 0 if they thought that none of the faces presented corresponded to the author of the crime. No response time limit was imposed.

PROCEDURE B was composed of two phases – the first one was a SEQL with Stopping Rule (SEQL-StopR), where the procedure was stopped after a positive identification was made. The lineup consisted of six faces, being either five distractors and the culprit seen in the video (in target present condition), or six distractors (in the target absent condition), which appeared sequentially until the first positive identification. The participants had to decide, for each face that was showed, if the individual they were seeing was the one they had seen in the video, pressing “S” if it was, or "N" if it was not. The presentation of the next image was not time-dependent, being triggered, or not, by their answer. Before each face, a fixation cross was shown for 500 ms. The second phase consisted of an SIML, equal to PROCEDURE A. No response time limit was imposed in both phases.

PROCEDURE C also consisted of two phases – the first one was a Passive Sequential Lineup (SEQL-Pass) composed by six faces, being five distractors and the culprit seen in the video (in the target present condition) or six distractors (in the target absent condition), which appeared sequentially, for 5000 ms each . As the name implies, participants simply looked passively at the faces, deciding nothing whatsoever. Before each face, a fixation cross was shown for 500 ms. The second phase consisted of an SIML, equal to PROCEDURE A. No response time limit was imposed in this phase.

Experimental tasks

Three experimental tasks (A, B and C) were built for this study, one for each Procedure (Figure 16), consisting of eight experimental blocks. Each block consisted of an instruction slide, followed by a theft video, a memory interference task and, finally, the corresponding procedure. As a memory interference task, participants had to perform a Classic Stroop Test (Colours – Computer Version), after which, according to their group (1, 2 or 3), they were exposed to the corresponding Procedure (Task A-Procedure A, Task B-Procedure B, Task C-Procedure C). SIML started with a drift correction (DC) point during 500 ms and, before each face in the SEQL, a fixation cross is shown for 500 ms.

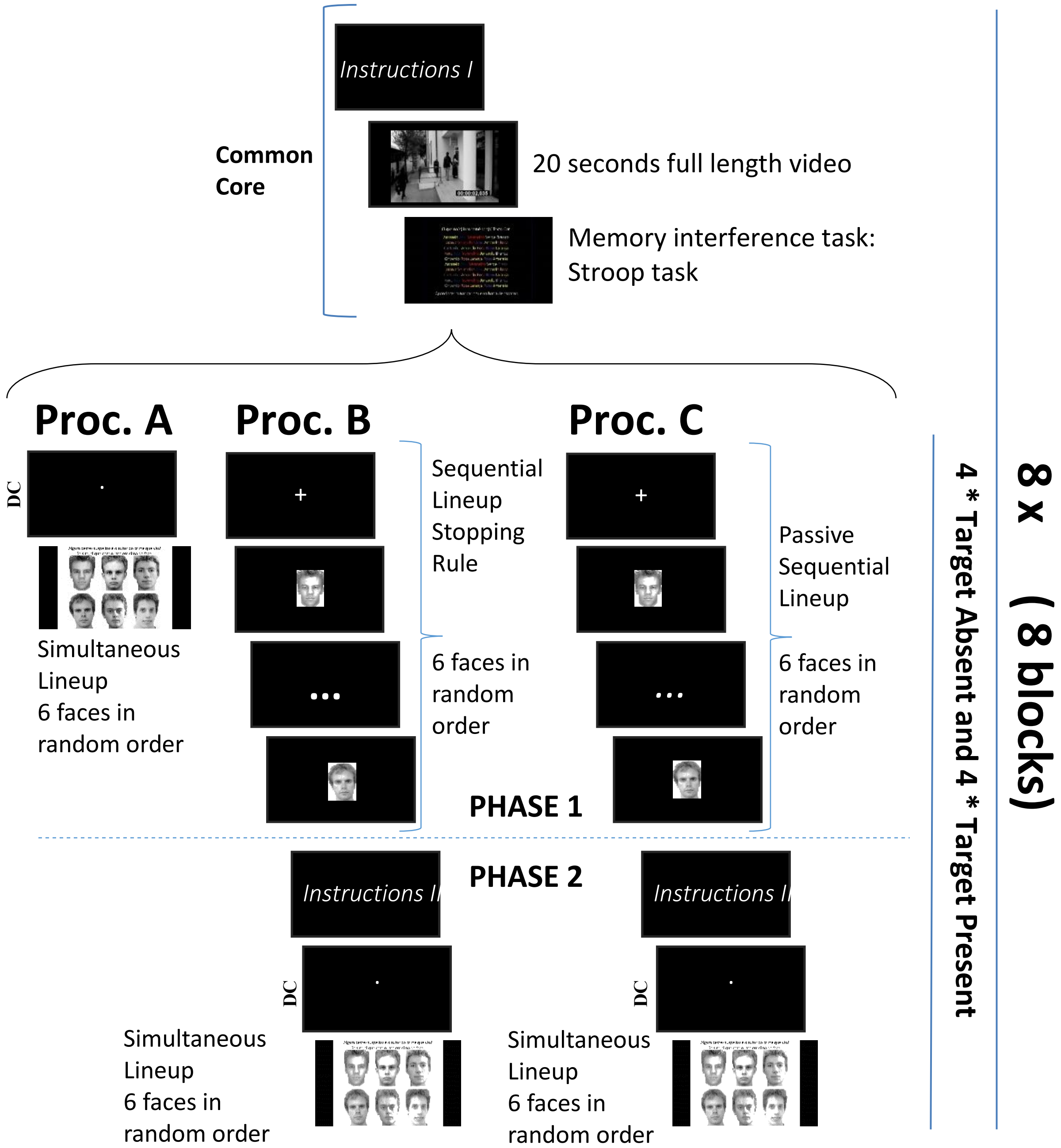


Figure 16. Tasks A, B, and C (Common Core + Correspondent Procedure)

Of the eight blocks that each participant saw in random order, four were target present, and four were target absent (the eight videos were counterbalanced). As can be seen from the description of the task, the instructions I of Procedure A were the same as the Instructions II of procedures B and C. The instructions I of procedure B were tailored for the SEQL-StopR. Instructions I for procedures A and B and instructions II for

Procedures B and C were made considering the "cautionary instructions" recommended by Malpass and Devine (1981). In these instructions, it is said that the lineup may not contain the perpetrator (see Clark, 2005). The Instructions I of Procedure C did not take into account these "cautionary instructions" because the participants did not have to make any decision, they were only invited to look carefully at the faces of the individuals who were sequentially appearing. The existence of two instructions (I and II) in procedures B and C was to convey the existence of two distinct tasks.

Equipment

The experimental task was displayed on a computer with a 19-inch monitor, using the Tobii Studio v3.0 *software*, while a Tobii T120 Eye Tracker was recording. The eyetracking system was calibrated using a 12-point task.

Design

The core experimental design was 3*2, where the type of procedure (A, B and C) was a between-subjects variable, and the target/culprit presence (Present or Absent) was a within-subjects variable. For procedure B the design was 3*2*2 with Phase (Phase 1 or Phase 2) as a within-subjects variable.

Dependent Measures

The behavioural measures were the percentage of each type of responses in each lineup using the Signal Detection Theory parameters. Thus, for each lineup procedure and phase, five different types of responses were registered: ***Hits*** (the participant identified the criminal), ***Misses*** (the participant identified nobody when the criminal was in the lineup), ***False Alarms*** *in target present* (the participant identified a distractor, when the target was in the lineup), ***False Alarms*** *in target absent* (the participant identified a distractor when the target was not present) and ***Correct Rejections*** (the participant did not identify anyone, and the target was not in the lineup).

The eyetracker's dependent variables were the *average dwell time in the target* (seven Regions of Interest (ROIs): total face, eyes, nose, mouth, hair, chin and ears), the *average dwell time in distractors* (seven ROIs: total face, eyes, nose, mouth, hair, chin, and ears) and the *average dwell time in the identified individual* (ID; not being the target - in case of a false alarm, in the same ROIs), using the same signal detection methodology. The *average number of visits/fixations* (number of fixations on a given face, after fixating another face) were analysed using same the procedure of dwell time but were only used in Phase 2 lineup comparisons, because in SEQLs we have a single photo show.

General Procedure

Upon arrival, participants were pseudo-randomly assigned (to ensure gender equity) to groups A, B or C. Subsequently, the participants signed the informed consent, and sat comfortably where the experiment was going to take place. Finally, the participants proceeded to the eyetracker calibration and performed the experimental task that corresponded to their group condition.

Data Analysis

For behavioral analysis, Kolmogorov Smirnov's tests show that only **False Alarms** in Phase 1 fitted with the curve of normal distribution. Some of the variables reached statistical normality using transformations, however using transformations in discrete variables is still not consensual and, hence, for the main behavioural analyses we decided to use non-parametric tests, specifically the U of Mann-Withney and Wilcoxon. However, in order to better explore the differences in the **Misses** and **False Alarms** in the target absent condition between groups 1 (Procedure A) and 2 (Procedure B), in the 1st phase, the variable "False_Alarms_1_Phase_Absent" was transformed using a square root function. Also, because the variable "**Misses**_1_Phase" was discrete with three categories and not normal, it was converted into a truncated and conditional continuous random variable (CRV; see Appendix), using a MATLAB algorithm, made purposely for the study. This algorithm was adapted from the method developed by Hasan, Rehman, and Bhatti (2016), maintaining the mean, variance and standard deviation of the discrete input variables and, then, transformed using a square root function. These variables were analysed with *t*-tests (paired and independent samples) and Hedges' *g*, and Cohen's *d* is provided as a measure of the effect size. The Cl effect size is also provided and indicates the chance that, for a randomly selected pair of individuals, the score of a person from the first group in each analysis is higher than the score of a person from the second group (Lakens, 2013). The ESCI software (Cumming, 2013) was used for replicability estimation. In these analyses (Criminal out – Miss and Innocent in – False Alarm) only the differences between Groups 1 and 2 in the first phase are explored since the results obtained with the parametric tests among the others were only for the number of **Hits** and **False Alarms**/Positives in the target present condition.

However, **False Alarms** have obtained the value 0 in the second group on 2nd phase and, therefore, cannot be transformed. Additionally, conversion to CRV does not work with pairwise designs.

Regarding dwell time and a number of fixations/visits, the data was 'square root' transformed before being submitted to an inferential statistical analysis. Because the pattern of results was the same for the transformed and untransformed scores, the descriptive statistics that are provided are based on the untransformed data. Again, these variables were analysed with *t*-tests, Cohen's *d* and CL effect sizes are also provided.

RESULTS

Behavioral Results

The average percentage of responses obtained in each category[20] - **Hits**, **Misses**, **False Alarms** Target Present, Target Absent **False Alarms** Target Absent and **Correct Rejections** - by each of the three groups are shown in Figure 17.

SIML (Group 1/Procedure A - Phase 1) vs. SEQL-StopR (Group 2/Procedure B - Phase 1)

Through visual data analysis, Figure 17 shows that we obtained more **Hits** in the SIML lineup than in the SEQL-StopR, but this difference was not statistically significant ($Mean\ Rank_{\text{SEQL-StopR}} = 18.30$; $Mean\ Rank_{\text{SIML}} = 22.70$; $U = 156.00$, $Z = -1.25$, $p = .210$). Still in the first phase, the percentage of **Misses** recorded by the group that performed the SIML (Procedure A) was significantly lower than the percentage obtained by the group that carried out the SEQL-StopR (Procedure B) ($Mean\ Rank_{\text{SEQL-StopR}} = 25.28$; $Mean\ Rank_{\text{SIML}} = 15.73$; $U = 104.50$, $Z = -2.84$, $p < .01$). In the **False Alarms** - Target Present the two groups again did not register statistically significant differences ($Mean\ Rank_{\text{SEQL-StopR}} = 18.38$; $Mean\ Rank_{\text{SIML}} = 22.63$; $U = 157.50$, $Z = -1.26$, $p = .208$). Regarding the **False Alarms** in the Target Absent, the group that saw the SIML obtained a higher percentage of **False Alarms** than the group that viewed the SEQL-StopR and this difference was statistically significant ($Mean\ Rank_{\text{SEQL-StopR}} = 16.55$; $Mean\ Rank_{\text{SIML}} = 24.45$; $U = 121.00$, $Z = -2.22$, $p < .05$). The group of participants who saw the SEQL-StopR obtained a

[20] As indicated before, due to the violation of the normality assumption, the statistical inference is made using non-parametric tests, and results in the text are given in Mean Ranks. However, in this figure, the averages of the percentages are presented to facilitate visual data analysis. It is relevant to point out that the pattern of effects was repeated with the use of parametric statistics.

significantly higher percentage of **Correct Rejections** than the group that performed the SIML (*Mean Rank* $_{\text{SEQL-StopR}}$ = 24.45; *Mean Rank* $_{\text{SIML}}$ = 16.55; U = 121.00, Z = -2.22, p < .05).

SEQL-StopR (Group 2/Procedure B - Phase 1) vs. SIML (Group 2/Procedure B - Phase 2)

In order to analyse the differences between group 2's responses (Procedure B) in the first and second phase, five *Wilcoxon Signed-Ranks Tests*[21] for paired samples were performed. Regarding the **Hits**, group 2 obtained significantly higher scores in phase 2 than in phase 1 (Z = -3.02, p < .01). Regarding **Misses**, although the scores were higher in phase 2, the differences were not statistically significant (Z = -.26, p = .797). For the **False Alarms** - Target Present, group 2 obtained significantly lower scores in Phase 2 - SIML (Z = -3.31, p < .001). For the **False Alarms** - Target Absent (Z = -2.66, p = .791) and **Correct Rejections** (Z = -2.66, p = .791), the scores obtained on the two phases were very similar and non-statistically significant.

All participants who performed a **False Alarm** (Target present) in the first phase (SEQL-StopR), hit the target (**Hits** Category) in the second phase, and only one participant moved from **Misses** to **Hits** from the first to the second phase.

SIML (Group 1/Procedure A - Phase 1) vs. SIML after SEQL-StopR (Group 2/Procedure B - Phase 2)

In this analysis, the comparison between the SIML without previous SEQL-StopR and the SIML after SEQL-StopR in the first phase was performed. Group 2, which saw SEQL before the SIML (Procedure B), obtained, in the second phase, a greater percentage of **Hits** than group 1 (*Mean Rank* $_{\text{SEQL-StopR}}$ = 24.38; *Mean Rank* $_{\text{SIML}}$ = 16.63; U

[21] Because the data were skewed in *Wilcoxon Signed-Ranks Test*, no ranks were displayed.

= 122.50, *Z* = -2.20, *p* < .05). Regarding the **Misses**, although group 1 had a lower percentage, this difference was only marginally significant (*Mean Rank* $_{SEQL\text{-}StopR}$ = 23.55; *Mean Rank* $_{SIML}$ = 17.45; *U* = 139.00, *Z* = -1.86, *p* = .063). The percentage of **False Alarms** - Target Present was significantly higher in group 1 (*Mean Rank* $_{SEQL\text{-}StopR}$ = 12.00; *Mean Rank* $_{SIML}$ = 29.00; *U* = 30.00, *Z* = -5.20, *p* < .001). On the other hand, the percentage of **False Alarms** - Target Absent, although also greater in group 1, the differences were not significant (*Mean Rank* $_{SEQL\text{-}StopR}$ = 17.45; *Mean Rank* $_{SIML}$ = 23.55; *U* = 139.00, *Z* = -1.69, *p* = .092). Respecting the **Correct Rejections**, a higher average value for group 2 was observed, but, again, the differences were not statistically significant (*Mean Rank* $_{SEQL\text{-}StopR}$ = 23.55; *Mean Rank* $_{SIML}$ = 17.45; *U* = 139.00, *Z* = -1.69, *p* = .092).

SIML after SEQL-StopR (Group 2/Procedure B - Phase 2) vs. SIML after SEQL-Pass (Group 3/Procedure C - Phase 2).

The percentage of **Hits** recorded by group 2 was significantly higher than that recorded by group 3 (*Mean Rank* $_{SEQL\text{-}StopR}$ = 24.45; *Mean Rank* $_{SEQL\text{-}Pass}$ = 16.55; *U* = 121.00, *Z* = - 2.22, *p* < .05). Inversely, with the **Misses** score which, although the **Misses** score was higher in the group that performed the SEQL-Pass first (group 3), differences did not reach statistical significance (*Mean Rank* $_{SEQL\text{-}StopR}$ = 19.33; *Mean Rank* $_{SEQL\text{-}Pass}$ = 21.68; *U* = 176.50, *Z* = -.68, *p* = .498). Regarding the **False Alarms** - Target Present, group 3 obtained significantly higher scores (*Mean Rank* $_{SEQL\text{-}StopR}$ = 16.00; *Mean Rank* $_{SEQL\text{-}Pass}$ = 25.00; *U* = 110.00, *Z* = -3.34, *p* < .01). The **False Alarms** - Target Absent (*Mean Rank* $_{SEQL\text{-}StopR}$ = 20.25; *Mean Rank* $_{SEQL\text{-}Pass}$ = 20.75; *U* = 195.00, *Z* = -.14, *p* = .890) and the **Correct Rejections** (*Mean Rank* $_{SEQL\text{-}StopR}$ = 20.75; *Mean Rank* $_{SEQL\text{-}Pass}$ = 20.25; *U* = 195.00, *Z* = -.14, *p* < .890) obtained exactly the same mean values in each group, and differences between groups were not significant.

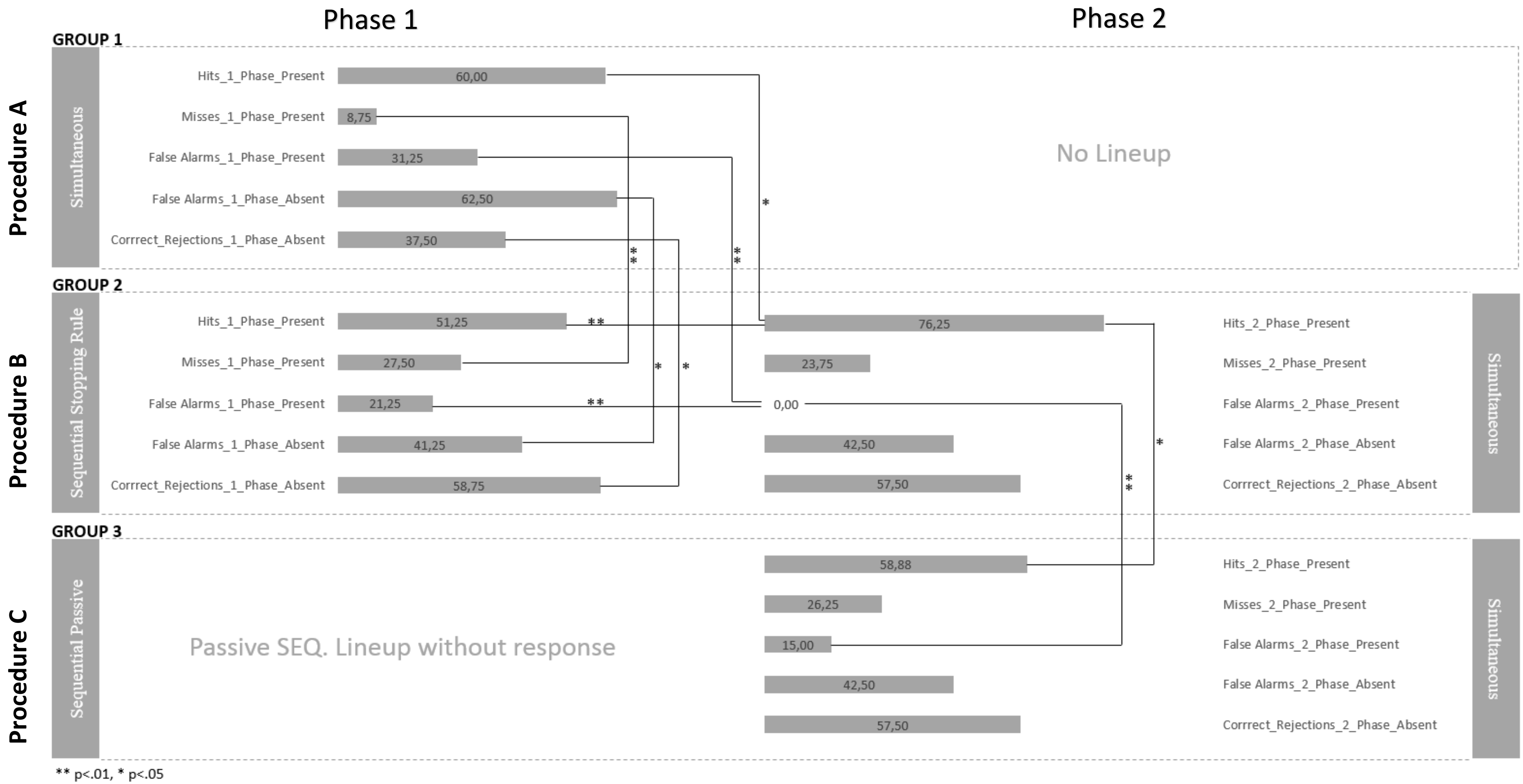


Figure 17. Illustration of multiple comparisons of percentage averages (Hits, Misses, False Alarms in Target Present, False Alarms in Target Absent and Correct Rejections) obtained in groups 1, 2 and 3 in the respective Procedure in phases 1 and 2.

(Criminal out) Miss and (Innocent in) False Alarm (SIML/Group 1 vs. SEQL-StopR/Group 2 - Phase 2)

As described in the Method section (data analysis), these analyzes were performed with transformed variables (CRV+Sqrt), and therefore they must be read with caution. Additionally, the descriptive statistics do not have an absolute interpretation and therefore are not presented.

Regarding the **Misses**, group 1 showed a smaller percentage than group 2.($t(38) = 3.90$, $p < .001$, Hedges' $g = 1.27$, Cohen's $d = 1.21$). The value of CL effect size (.81) indicates that, if we randomly select a pair of individuals (one from each group), the probability that the score of the participant of group 1 is lower than that of the participant of group 2 is about 81%. Also, approximately 88.8% of the scores distribution for the **Misses** of the participants that viewed the SEQL-StopR lineup was above the average of the scores of the individuals who visualized the SIML lineup. From a replicability perspective, the ESCI software (Cumming, 2013) was used, where 25 experiments using the data from our study were simulated. The results are shown in Figure 18.

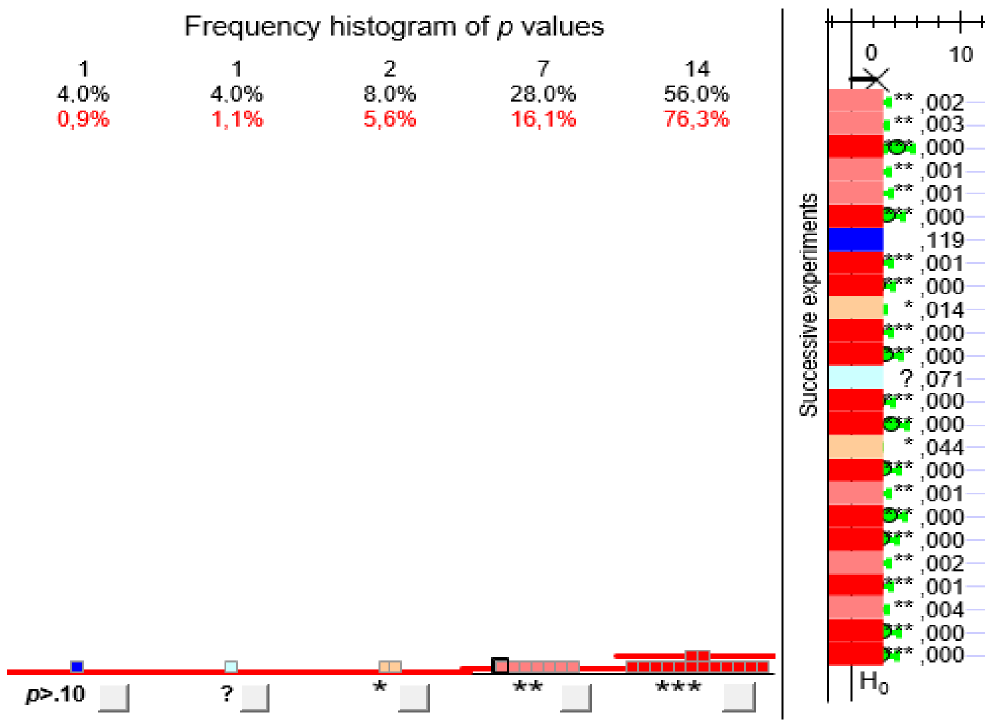


Figure 18. Histogram of frequencies of p values from the replication of twenty-five experiments testing the differences in the Miss responses (black values represent the observed frequencies; red values represent the expected frequencies).

By analyzing Figure 18, we can verify that, in the 25 simulated successive replications, only two of them were not statistically significant and one of them reached a statistical trend. The expected proportions for the two no-effects categories ($p > .10$ and $p > .05$) were higher in observed replications than in the expected.

Regarding the **False Alarms - Target Absent** in the first phase, group 1 that visualized the SIML had a higher percentage of FA than group 2 that visualized a SEQL ($t(38) = 2.48$, $p < .05$, Hedges' $g = .77$, Cohen's $d = .80$). The value of CL effect size (.71) indicates that, if we randomly select a pair of individuals (one from each group), there is a 71% of probability that the score of the member of group 1 is lower. About 74.6% of the distribution of Target Absent False Alarm scores for group 1 individuals was above the average scores of individuals of group 2. Again, 25 experiments were simulated using the data from our study. The results are shown in Figure 19.

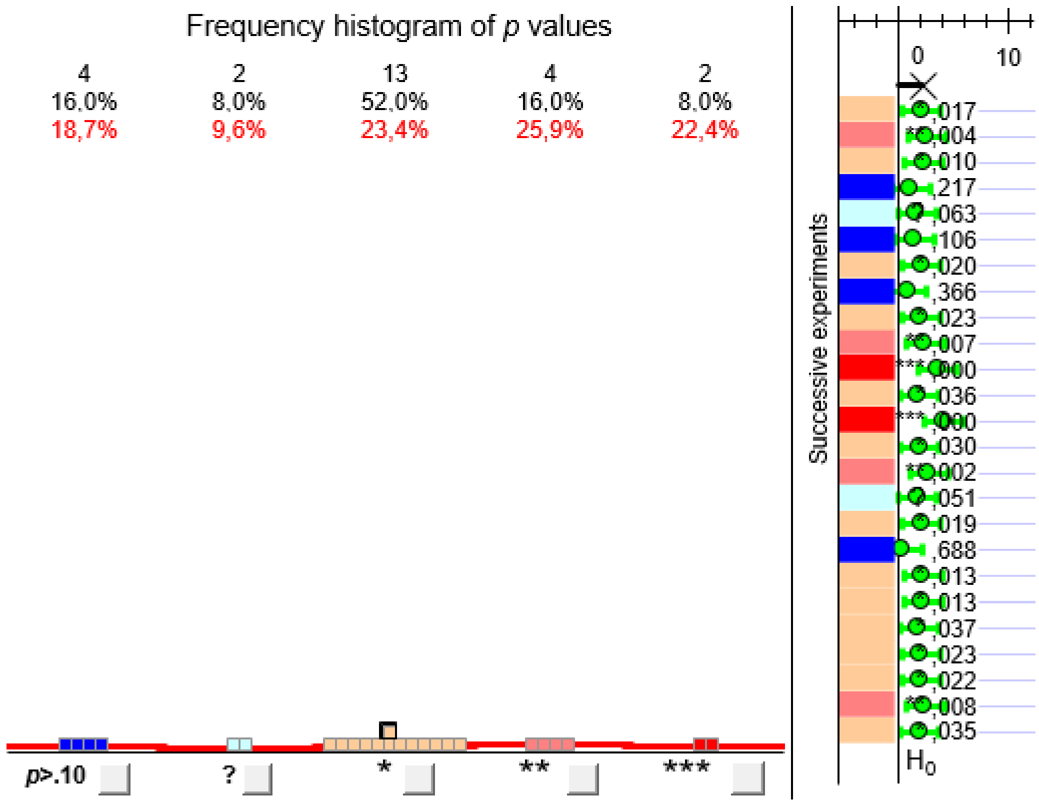


Figure 19. Histogram of frequencies of p values from the replication of twenty-five experiments testing the differences in the False Alarm – Target Absent responses (black values represent the observed frequencies; red values represent the expected frequencies).

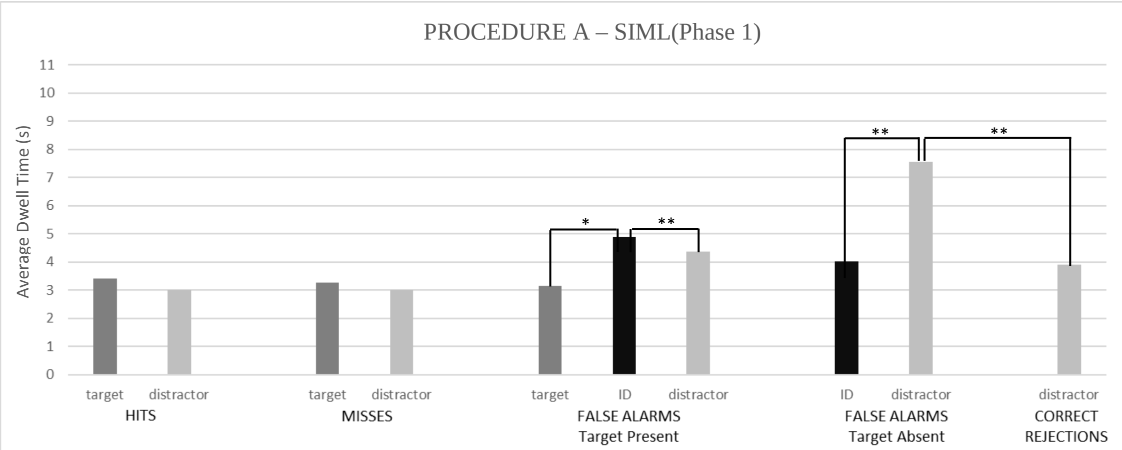
Considering our data, larger proportions of cases where the null hypothesis is not violated were expected. Nevertheless, the proportion of replications in which the *p* was below .01 was lower than expected.

Gazze Patterns

Average dwell time length in all Face

Average dwell time on Face - Phase 1 – Procedure A - SIML vs. Phase 1 – Procedure B - SEQL-StopR

The average dwell time on targets, distractors and identified individual (ID; non-target), in each response category, divided by procedure, is shown in Figure 20.



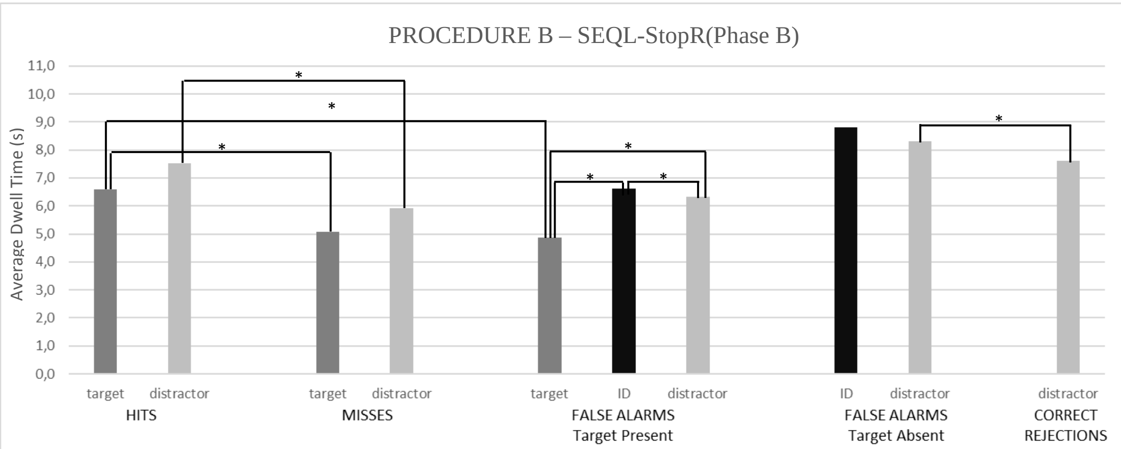


Figure 20. Average dwell time for each response category for targets, distractors and Identified Individual (ID), by Procedure (Phase 1 – Procedure A - SIML vs. Phase 1 – Procedure B - SEQL-StopR).
P<.05 *p<.01

In relation to the **Hits**, a mixed ANOVA, 2 within (type - target vs. distractor) x 2 between (procedure - SIML vs. SEQL-StopR), showed a main effect of procedure ($F(1,35) = 29.76$, $p < .001$, $\eta_p^2 = .46$), but no effect of type ($F(1,35) = .19$, $p = .667$, $\eta_p^2 = .005$), nor of interaction ($F(1,35) = 1.79$, $p = .190$, $\eta_p^2 = .05$). Splitting the ANOVA by Procedure neither of the two target - distractor comparisons were significant.

In the **Misses**, the mixed ANOVA, 2 within (type - target vs. distractor) x 2 between (procedure - SIML vs. SEQL-StopR), showed a main effect of procedure ($F(1,19) = 17.27$, $p < .001$, $\eta_p^2 = .48$), but no effect of type ($F(1,19) = .30$, $p = .590$, $\eta_p^2 = .02$) nor of interaction ($F(1,19) = 1.98$, $p = .175$, $\eta_p^2 = .01$) were verified. As for the **Hits**, we did a split between the two procedures but neither of the two target - distractor comparisons proved to be significant.

In order to analyze the **False Alarms - Target Present**, a mixed ANOVA, 3 within (type - target vs. identified vs. distractor) x 2 between (procedure - SIML vs. SEQL-StopR), was performed. The data showed absence of interaction ($F(2,54) = .04$, $p = .109$, $\eta_p^2 = .004$), but showed main effects of procedure ($F(1,27) = 5.88$, $p <.05$, $\eta_p^2 = .18$) and type ($F(2,54) = 9.08$, $p < .01$, $\eta_p^2 = .26$). For procedure A, comparisons showed that the differences recorded between the average dwell time at the face of the identified individual and the average dwell time looking at the face of the target (see Figure 20 - top panel) assumed statistical significance ($t(16) = 2.20$, $p < .05$, Cohen's $d = 0.55$, CL effect size = .71). The same happened with the difference between the average dwell time on the face of the identified individual and the average dwell time at the face of the distractors ($t(16) = 3.00$, $p < .01$, Cohen's $d = 0.75$, CL effect size = .77)[22]. The

[22] Although there is a greater difference in absolute value in the target - ID comparison, this comparison shows higher inferential results due to the very small (.03) standard error of the mean.

difference between the average dwell time at the face of the distractor and the average dwell time looking at the face of the target only reached a statistical trend ($t(16) = 1.91$, $p = .075$, Cohen's $d = 0.48$, CL Effect size = .68). Regarding procedure B, the absolute differences observed in Figure 16 (down panel) between the average dwell time at the face of the identified individual and the average dwell time at the face of the target were statistically significant ($t(12) = 2.97$, $p < .05$, Cohen's $d = 0.82$, CL effect size = .80). The same happened with the difference between the average dwell time at the face of the identified individual and the average dwell time looking at the face of the distractor ($t(12) = 2.86$, $p < .05$, Cohen's $d = 0.79$, CL effect size = .79). Also, the difference between the average dwell time looking at the face of the distractor and the average dwell time looking at the face of the target obtained statistical meaning ($t(12) = 2.33$, $p < .05$, Cohen's $d = 0.65$, CL effect size = .74).

With respect to the **False Alarms - Target Absent** the mixed ANOVA, 2 within (type – identified individual vs. distractor) x 2 between (procedure - SIML vs. SEQL-StopR), showed a main effect of type ($F(1,36) = 194.06$, $p < .001$, $\eta_p^2 = .84$), procedure ($F(1,36) = 9.85$, $p < .01$, $\eta_p^2 = .22$) and of interaction ($F(1,36) = 296.74$, $p < .001$, $\eta_p^2 = .89$). In relation to procedure A, the absolute differences between the average dwell time looking at the face of the identified individual and average dwell time looking at the face of the distractor assumed statistical significance ($t(18) = 27.71$, $p < .001$, Cohen's $d = 6.36$, CL effect size = 1). Regarding procedure B, the differences between the average dwell time looking at the face of the identified individual and the average dwell time looking at the distractors' face revealed a statistical trend ($t(18) = 1.99$, $p = .062$, Cohen's $d = 0.67$, CL effect size = .67). Importantly, for this category of response (**False Alarms** -

Target Absent), there was no significant difference between the average dwell time looking at the face of the distractor between the two procedures.

Within lineup, procedure comparisons were made between categories of both **Target Present** and between categories of the **Target Absent**. The results for procedure A showed that there were only differences between categories in the **Target Absent**. Actually, when participants made a **False Alarm**, the average dwell time on the distractor when participants made a false alarm was significantly greater then when they correctly rejected the lineup ($t(13) = 5.17$, $p < .001$, Cohen's $d = 1.38$, CL effect size = .91). Regarding procedure B, statistically significant differences were obtained between the categories of the **Target Present** and the **Target Absent** conditions. In fact, the average dwell time on the targets was significantly higher for the **Hits** than for the **Misses** ($t(13) = 2.85$, $p < .05$, Cohen's $d = 0.76$, CL effect size = .77) and for the **False Alarms** ($t(12) = 2.58$, p < .05, Cohen's $d = 0.72$, CL effect size = .76). In relation to distractors, this difference only occurred between **Hits** and **Misses**, with participants dwelling longer on distractors when they made a Hit than when they made a **Miss** ($t(12) = 2.79$, $p < .05$, Cohen's $d = 0.78$, CL effect size = .78). Regarding the **Target Absent** condition, when the participants make a **False Alarm** the average dwell time on the distractor is significantly greater than when they correctly reject the lineup ($t(17) = 2.32$, $p < .05$, Cohen's $d = .64$, CL effect size = .74).

Average dwell time on Face – Phase 1 - Procedure A - SIML vs. Phase 2 – Procedure B – SIML.

The average dwell time on targets, distractors and identified individual (non-target), in each response category, per procedure, is shown in Figure 21.

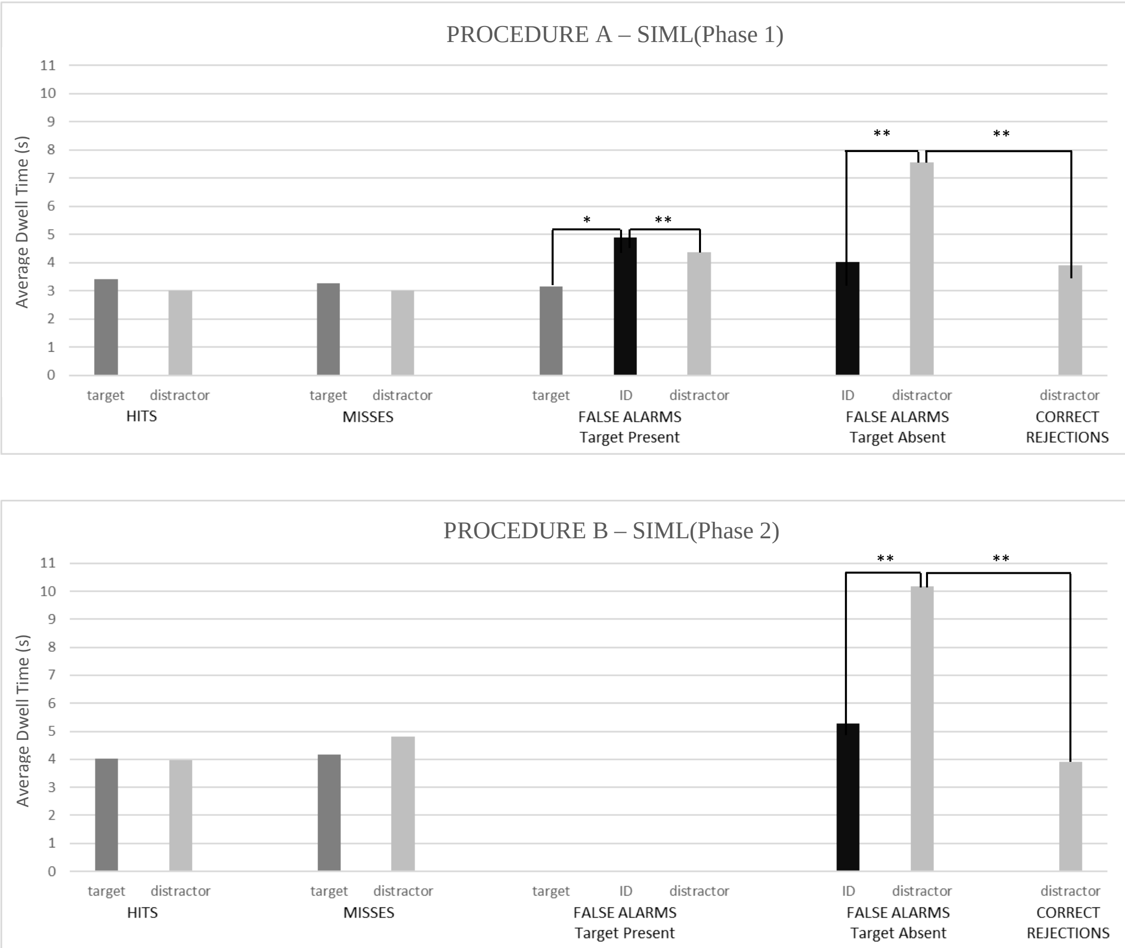


Figure 21. Average dwell time for each response category for targets, distractors and Identified Individual ID, by Procedure (Phase 1 - Procedure A - SIML vs. Phase 2 – Procedure B – SIML).

Relatively to the **Hits,** the results of the mixed ANOVA, 2 within (type - target vs. distractor) x 2 between (procedure - SIML A vs. SIML B), showed the lack of a main effect of procedure ($F(1,37) = 1.34$, $p = .254$, $\eta_p^2 = .04$), type ($F(1,37) = .42$, $p = .522$, $\eta_p^2 = .01$), and of interaction ($F(1,37) = .26$, $p = .614$, $\eta_p^2 = .01$).

In the **Misses**, the mixed ANOVA, 2 within (type - target vs. distractor) x 2 between (procedure - SIML A vs. SIML B), verified the absence of a main effect of type ($F(1,15) = .24$, $p = .633$, $\eta_p^2 = .02$) and of interaction ($F(1,15) = 2.12$, $p = .166$, $\eta_p^2 = .12$). However, a main effect of procedure ($F(1,15) = 4.86$, $p = .042$, $\eta_p^2 = .26$) was found. Relevantly, for this category of response (**Misses**), there is a significant difference between the average dwell time looking at the face of the distractor between the two procedures ($t(15) = 2.68$, $p < .05$, Cohen's $d = 1.20$, CL effect size = .81).

Due to the absence of **False Alarms** - Target Present in Procedure B – Phase 2, no ANOVA was carried out.

Regarding the **False Alarms - Target Absent,** a mixed ANOVA 2 within (type - identified vs. distractor) x 2 between (procedure - SIML A vs. SIML B), showed an effect of type ($F(1,29) = 652.135$, $p < .001$, $\eta_p^2 = .957$), but no effect of procedure ($F(1,29)$ = 2,937, $p = .097$, $\eta_p^2 = .092$) nor interaction, which only achieved a statistical trend ($F(1,29) = 3.598$, $p = .069$, $\eta_p^2 = .110$). As in procedure A (see comparison above), in procedure B the average dwell time on the face was significantly higher for the distractors ($t(11) = 6.13$, $p < .001$, Cohen's $d = 1.85$, CL effect size = .97).

Within lineup, procedure comparisons were made between categories of the **Target Present** and of the **Target Absent** conditions. Regarding procedure B, the average dwell time on the distractors when the participants make a false alarm is significantly greater than when they correctly reject the lineup ($t(7) = 3.90$, $p < .01$, Cohen's $d = 1.47$, CL effect size = .93).

Average dwell time on Face – Phase 1 - Procedure B - SEQL-StopR vs. Phase 2 - Procedure B – SIML.

The average dwell time on targets, distractors and identified Individual, in each response category, per phase of procedure B, is shown in Figure 22.

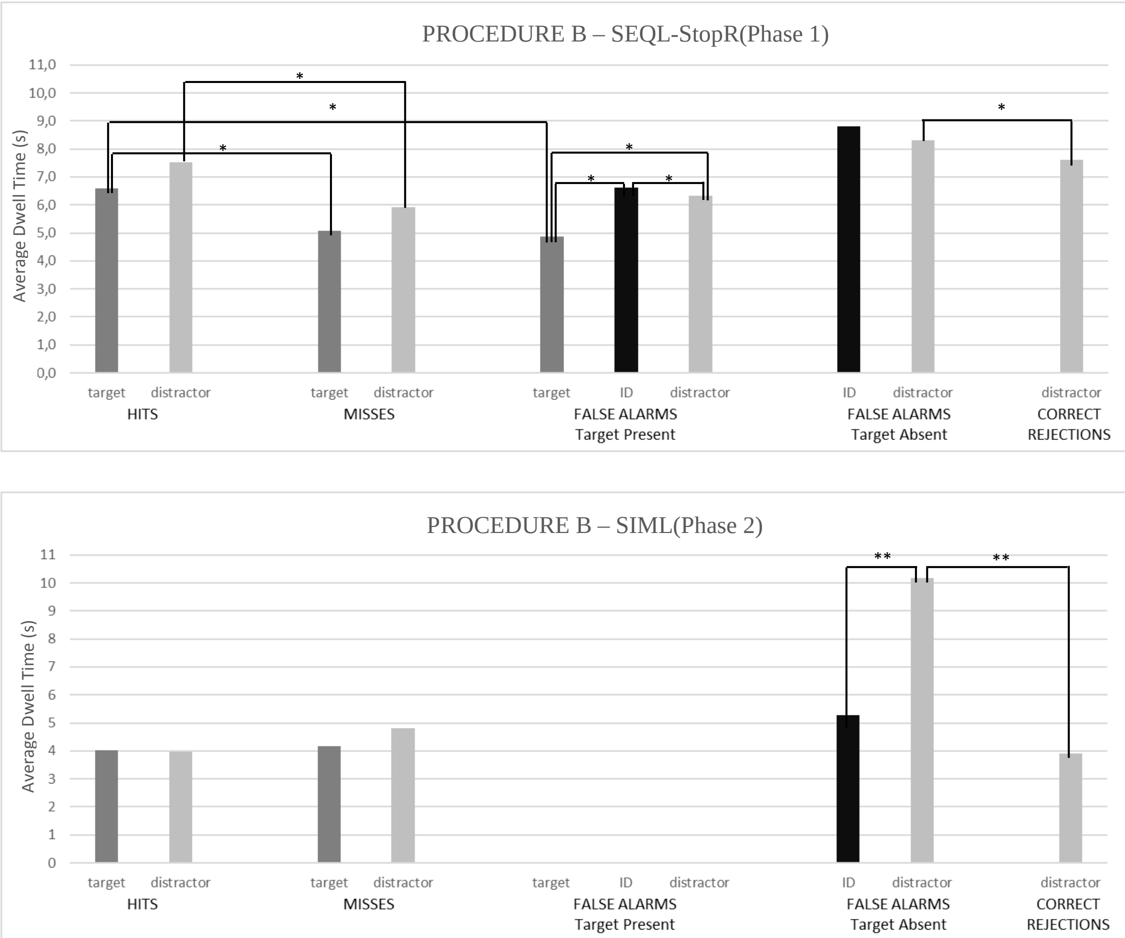


Figure 22. Average dwell time for each response category for targets, distractors and Identified Individual ID, by Procedure (Phase 1 - Procedure B - SEQL-StopR vs. Phase 2 - Procedure B – SIML).

Relatively to the **Hits,** the results of the repeated measures ANOVA, 2 within (type - target vs. distractor) x 2 within (phase/line up - SEQL-StopR B vs. SIML B), showed a main effect of phase/lineup ($F(1,17) = 11.51$, $p < .01$, $\eta_p^2 = .40$), but there was no effect of type ($F(1,17) = .80$, $p = .38$, $\eta_p^2 = .05$), nor of interaction ($F(1,17) = .56$, $p = .46$, $\eta_p^2 = .03$).

In the **Misses**, the repeated measures ANOVA, full within, verified the absence of main effect of type ($F(1,7) = 1.87$, $p = .214$, $\eta_p^2 = .21$) and a statistical trend in the interaction ($F(1,13) = 4.89$, $p = .063$, $\eta_p^2 = .41$). Although there was no main effect of phase/lineup ($F(1,7) = 5.47$, $p = .052$, $\eta_p^2 = .44$), the level of significance was marginal.

Given the absence of **False Alarms** - Target Present in Procedure B – Phase 2, the ANOVA was not produced.

With regard to the **False Alarms - Target Absent,** the mixed ANOVA, 2 within (type - identified vs. distractor) x 2 between (procedure - SIML A vs. SIML B), showed an effect of type ($F(1,30) = 84.62$, $p < .001$, $\eta_p^2 = .89$), and an interaction ($F(1,30) = 144.60$, $p < .001$, $\eta_p^2 = .93$). No effect of phase/lineup was observed ($F(1,30) = .62$, $p = .449$, $\eta_p^2 = .05$).

Average dwell time on Face – Procedure B - Phase 2 - SIML vs. Procedure C - Phase 2 - SIML (After SEQL-Pass).

The average dwell time on targets, distractors and identified Individual, in each response category, per procedure, is shown in Figure 23.

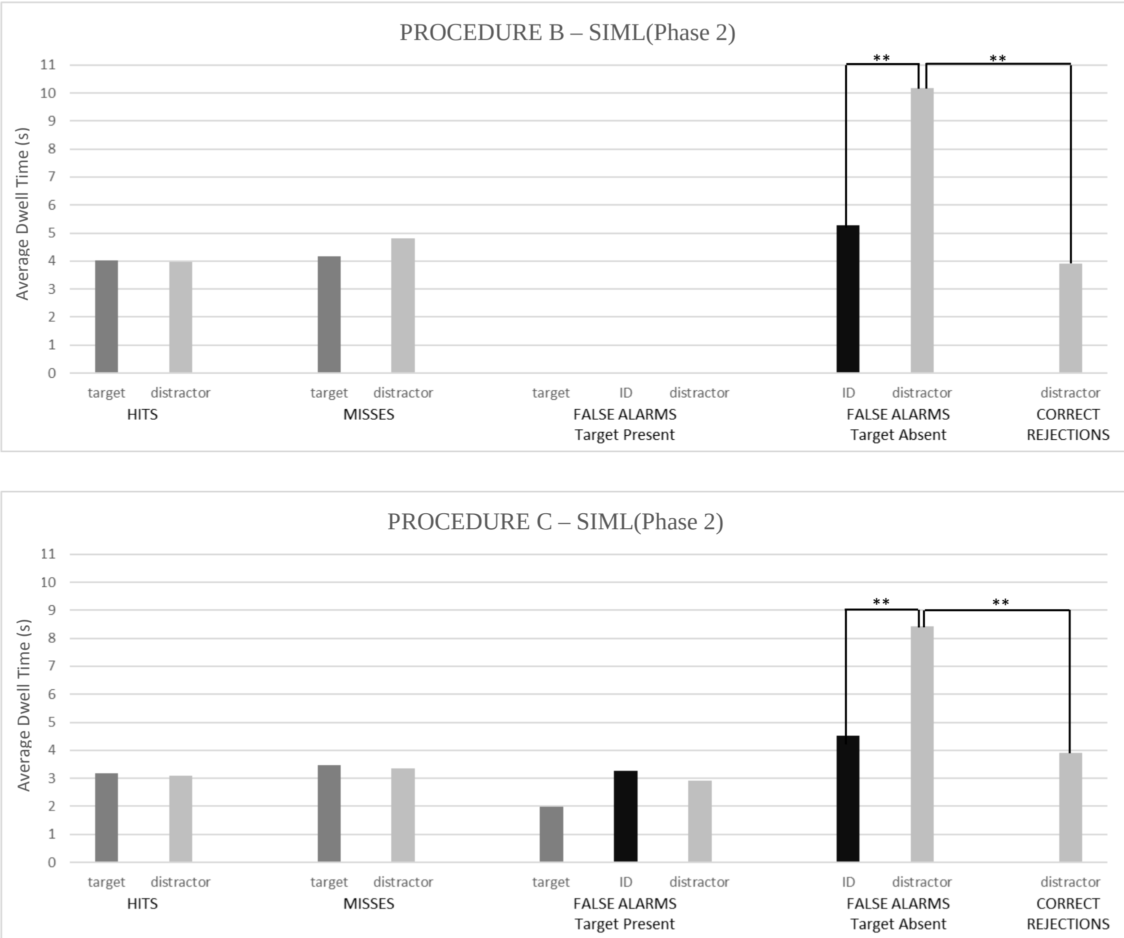


Figure 23. Average dwell time for each response category for targets, distractors and Identified Individual ID, by Procedure (Phase 2 - Procedure B - SIML vs. Phase 2 - Procedure C – SIML).

In relation to **Hits,** a mixed ANOVA, 2 within (type - target vs. distractor) x 2 between (procedure - SIML B vs. SIML C), showed the lack of a main effect of procedure (F(1,38) = 1.82, p = .186, η_p^2 = .05), effect of type (F(1,38) = .002, p = .966, η_p^2 = .000), and of interaction (F(1,38) = .01, p = .924, η_p^2 = .00).

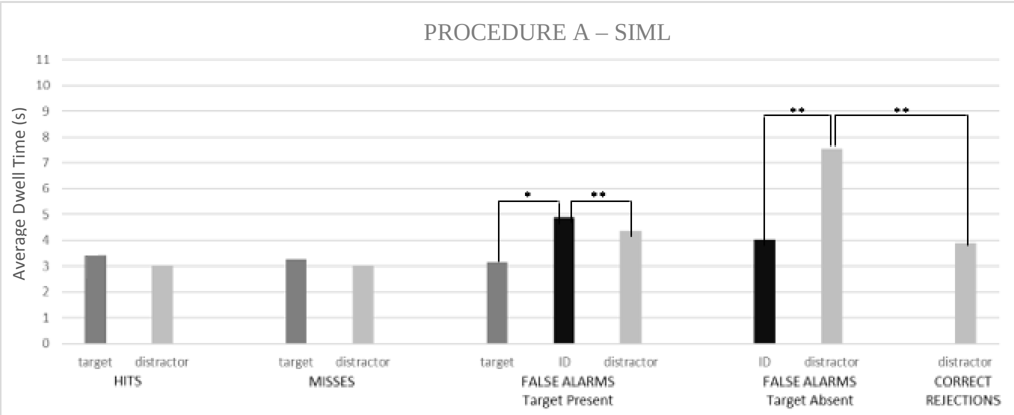
Regarding the **Misses**, the mixed ANOVA, 2 within (type - target vs. distractor) x 2 between (procedure - SIML B vs. SIML C), showed no effects of type ($F(1,24) = .04$, p =. 84, $\eta_p^2 = .002$), procedure ($F(1,24) = 2.94$, $p = .099$, $\eta_p^2 = .11$), or interaction ($F(1,24)$ = 2,19, $p = .152$, $\eta_p^2 = .08$).

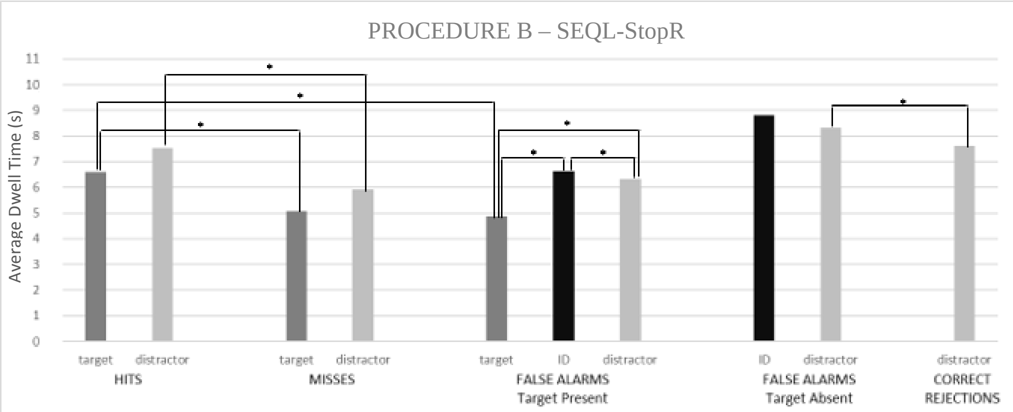
Given the absence of **False Alarms - Target Present** for one of the types of procedure, the ANOVA was not produced.

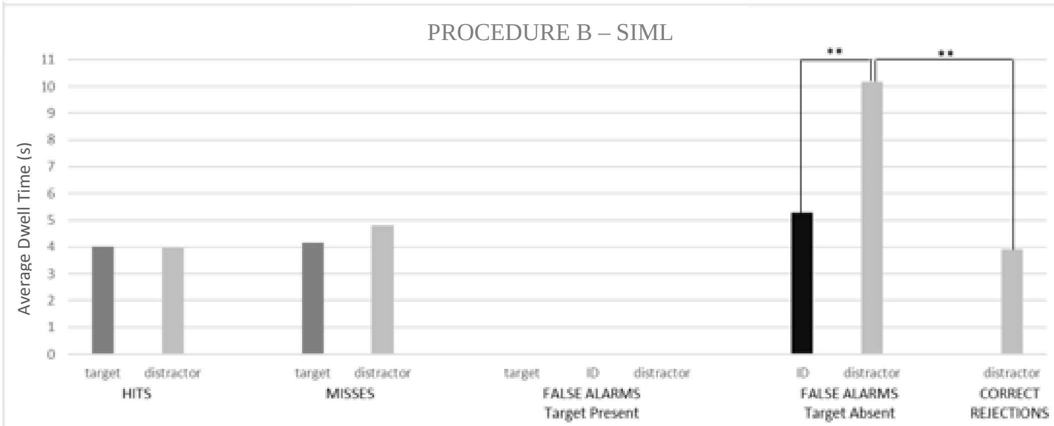
Concerning the **False Alarms - Target Absent,** a mixed ANOVA, 2 within (type – identified vs. distractor) x 2 between (procedure - SIML B vs. SIML C), showed an effect of type ($F(1,26) = 468.41$, $p < .001$, $\eta_p^2 = .95$), but not of procedure ($F(1,26) = .92$, p = .346, $\eta_p^2 = .03$) nor an interaction ($F(1,26) = 1.73$, $p = .200$, $\eta_p^2 = .06$). As in procedure A, in procedure C the average dwell time was significantly greater for the distractors than for the identified individual ($t(15) = 10.64$, $p < .001$, Cohen's $d = 2.65$, CL effect size = .99).

Within lineup, procedure comparisons were made between categories of the **Target Present** and the **Target Absent** conditions. Regarding procedure C, the average dwell time on the distractors when the participants made a false alarm was significantly higher than when they correctly rejected the lineup ($t(13) = 4.77$, $p < .001$, Cohen's d = 1.27, CL effect size = .89).

Due to the complexity of the results, Figure 24 shows an overall representation of the results in the various conditions.







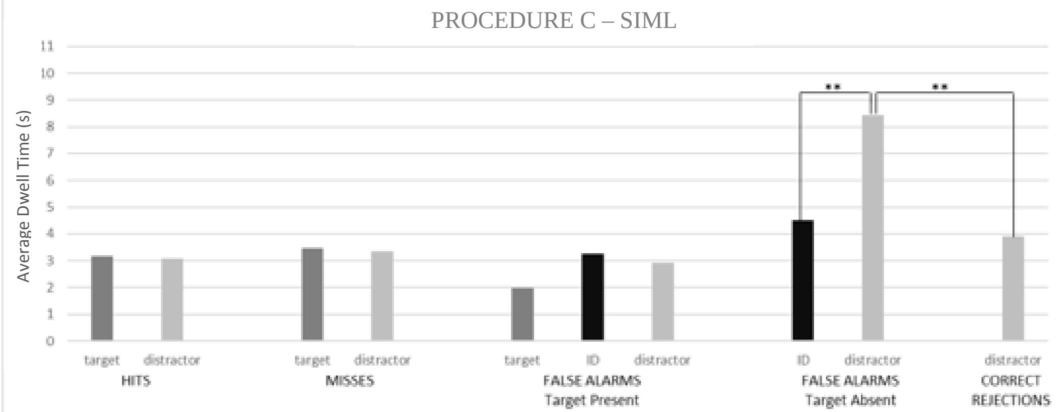


Figure 24. Overview of Dwell time Results for all the Procedures

Number of fixations/visits in all face

Average number of fixations in the faces – Phase 1, Procedure A - SIML vs. Phase 2, Procedure B - SEQL-StopR

The average number of fixations in the targets, distractors and identified individuals, in each response category, separated by procedure, is shown in Figure 25.

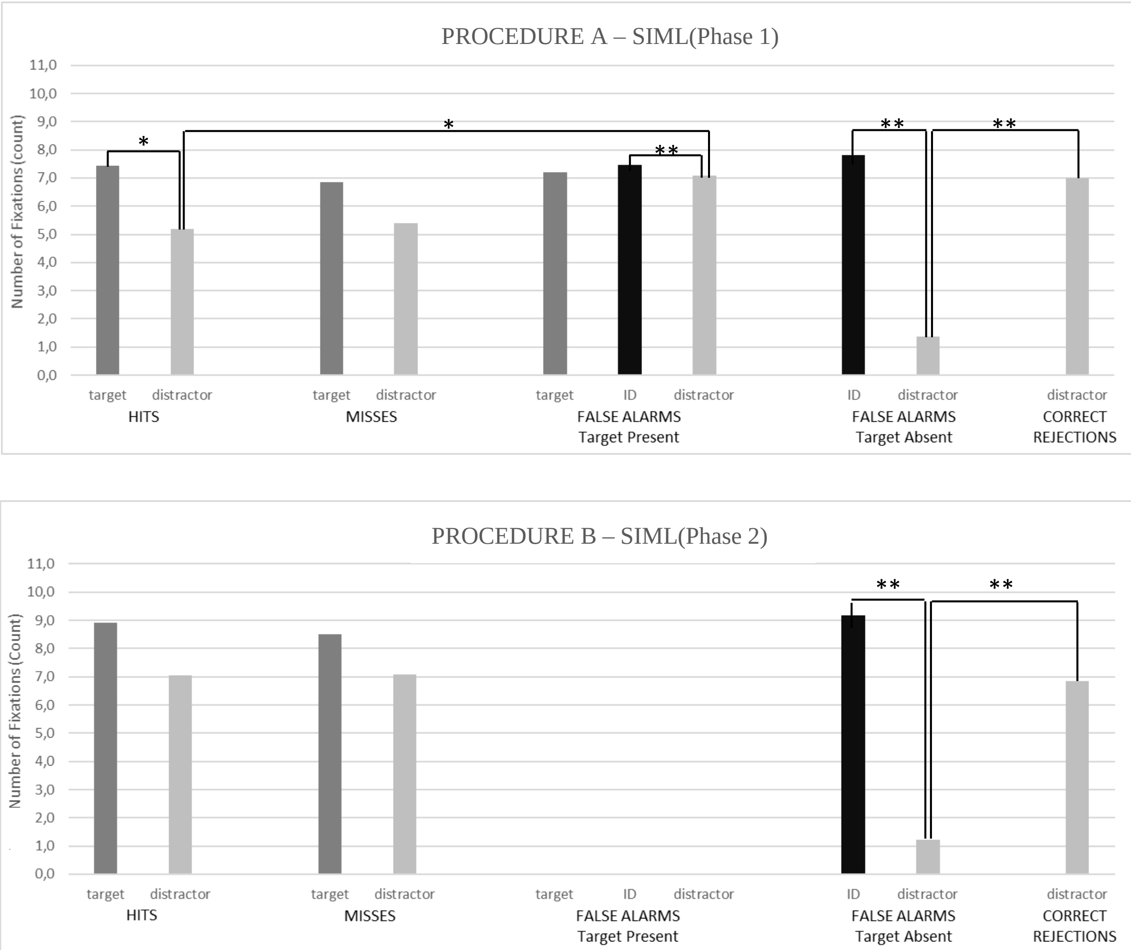


Figure 25. Number of fixations for each response category for targets, distractors and Identified Individual ID, by Procedure (Phase 1 – Procedure A - SIML vs. Phase 2 – Procedure B – SIML after SEQL-StopR).

For the **Hits,** the results of the mixed ANOVA, 2 within (type - target vs. distractor) x 2 between (procedure - SIML A SIM versus SIML B SIM), showed an effect of stimulus type (F(1,37) = 8.69 p < .001, η_p^2 = .19), but no effect of procedure (F(1,37) = 1.78, p = .191, η_p^2 = .05) nor interaction (F(1,37) = .08, p = .776, η_p^2 = .002). Only for procedure A, the number of visits to the target face was significantly higher than to the distractor (t(18) = 2.42, p < .05, Cohen's d = .55, CL effect size = .71).

In the **Misses**, the mixed ANOVA, 2 within (type - target vs. distractor) x 2 between (procedure - SIML A vs. SIML B), verified the absence of main effect of type (F(1,15) = 2.71, p = .120, η_p^2 = .15), procedure (F(1,15) = 2.73, p = .120, η_p^2 = .15) and of interaction (F(1,15) = .00, p = .999, η_p^2 = .00). Pairwise comparisons showed that only the difference between the identified individual and the distractor reached statistical significance (t(15) = 17.30, p < .001, Cohen's d = 4.33, CL effect size = .99).

As for the **False Alarms - Target Absent,** the mixed ANOVA, 2 within (type - identified vs. distractor) x 2 between (procedure - SIML A vs. SIML B), showed an effect of type (F(1,29) = 438,629, p < .001, η_p^2 = .938) and an interaction (F(1,29) = 5.16, p < .05, η_p^2 = .15). However, no effect of procedure (F(1,29) = .22, p = .641, η_p^2 = .01) was obtained.

Regarding procedure A, the average number of fixations on the face of the identified individual was significantly greater than the average number of fixations on the distractor faces (t(18) = 27.65, p < .001, Cohen's d = 6.52, CL effect Size = 1). The same is true for procedure B (t(11) = 10.01, p < .001, Cohen's d = 2.89, CL effect size = .99).

Within lineup, procedure comparisons were made between categories of the **Target Present** and the **Target Absent** conditions. Regarding procedure A, the average number of fixations on the distractor, when **Hits** occurred, was significantly higher than when **False Alarms** in **Target Absent** situations happened (t(14) = 2.19, p < .05, Cohen's d = .56, CL effect size = .71). Also, the average number of fixations on the distractor faces when **Correct Rejections** took place was significantly higher than than when **False**

Alarms in **Target Absent** situations occurred (*t*(12) = 14.15, *p* < .001, Cohen's *d* = 3.93, CL effect size = .99).

Regarding procedure B, the average number of fixations on the distractors when participants correctly rejected the lineup was significantly greater than when they made a false alarm (*t*(7) = 6.45, *p* < .001, Cohen's *d* = 1.47, CL effect size = .93).

Average number of fixations in the face – Phase 2, Procedure B - SIML vs. Phase 2, Procedure C – SEQL-Pass

The average number of fixations on the targets, distractors and identified individuals, in each response category, per procedure, is shown in Figure 26.

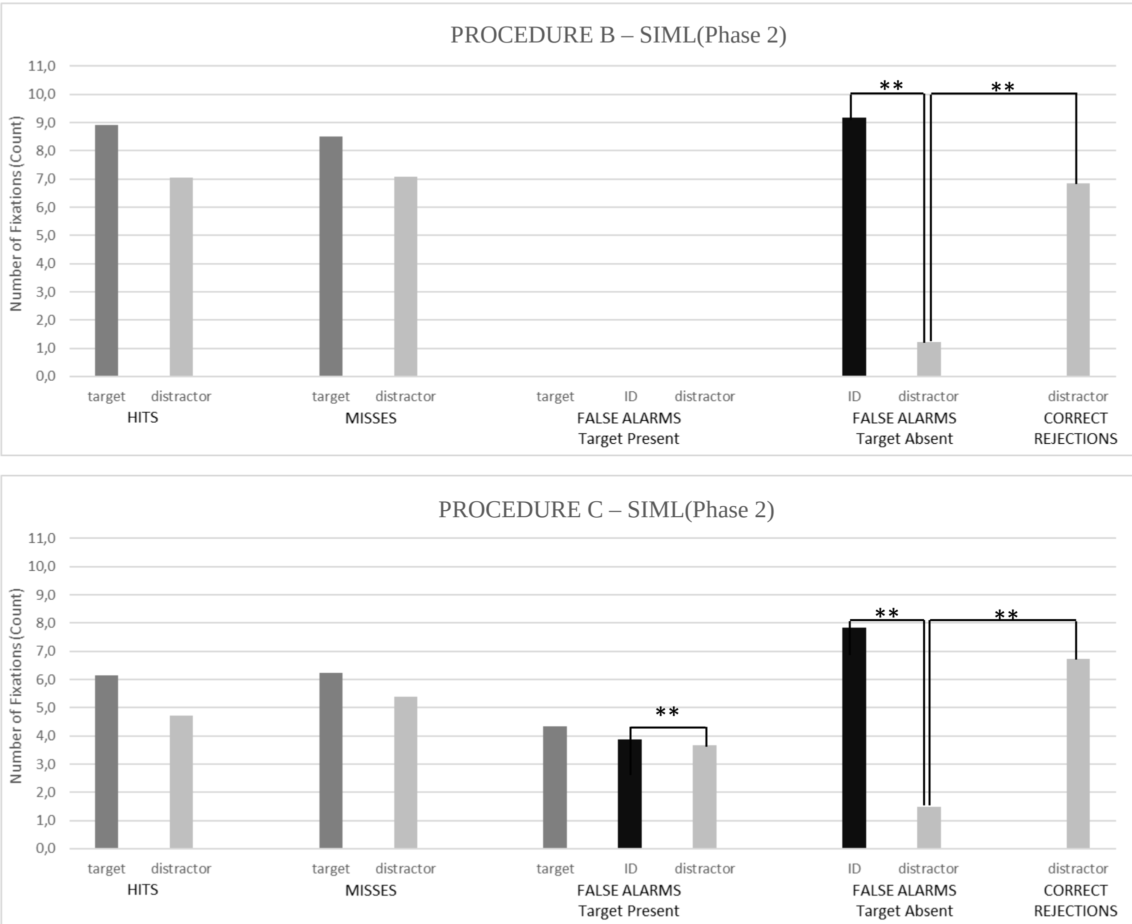


Figure 26. Number of fixations for each response category for targets, distractors and Identified Individual ID, by Procedure (Phase 2, Procedure B - SIML vs. Phase 2, Procedure C – SEQL-Pass).

In relation to the **Hits,** the results of the mixed ANOVA, 2 within (type - target vs. distractor) x 2 between (procedure - SIML Bvs. SIML C), showed an effect of type ($F(1,38) = 5.80$, $p < .05$, $\eta_p^2 = .13$) and of procedure ($F(1,38) = 5.02$, $p < .05$, $\eta_p^2 = .12$), but there was no interaction ($F(1,38) = .49$, $p = .83$, $\eta_p^2 = .001$). No significant differences between targets and distractors were observed in any of the procedures.

In the **Misses**, the mixed ANOVA 2 within (type - target vs. distractor) x 2 between (procedure - SIML B vs. SIML C), verified a statistical trend for the effect of type ($F(1,24) = 3.56$, $p = .071$, $\eta_p^2 = .13$) and procedure ($F(1,19) = 3.20$, $p = .09$, $\eta_p^2 = .18$), but no interaction was found ($F(1,19) = .03$, $p = .868$, $\eta_p^2 = .001$).

Considering the absence of **False Alarms - Target Present**, the ANOVA was not produced. For procedure C, the comparisons showed that only the difference between the identified individual and the distractor reached statistical significance ($t(8) = 10.92$, $p < .001$, Cohen's $d = 3.64$, CL effect size = .99).

Respecting the **False Alarms - Target Absent** the mixed ANOVA, 2 within (type - identified vs. distractor) x 2 between (procedure - SIML B vs. SIML C), showed an effect of type ($F(1,26) = 269.70$, $p < .001$, $\eta_p^2 = .91$), yet there was no effect of procedure ($F(1,26) = .03$, $p = .860$, $\eta_p^2 = .001$), nor interaction ($F(1,26) = 2.85$, $p = .103$, $\eta_p^2 = .099$).

Regarding procedure C, the total number of fixations on the face of the identified individual was significantly greater than the number of fixations on the distractor faces ($t(15) = 13.43$, $p <.001$, Cohen's $d = 3.36$, CL effect Size = .99). Within lineup, procedure comparisons were made between the categories of the **Target Present** and the **Target Absent** conditions. In procedure C, the average number of fixations on the distractors when participants correctly rejected the lineup was significantly greater than when they made a false alarm ($t(13) = 5.84$, $p < .001$, Cohen's $d = 1.56$, CL effect size =. 94).

Due to the complexity of the results, Figure 27 shows an overall picture of the results.

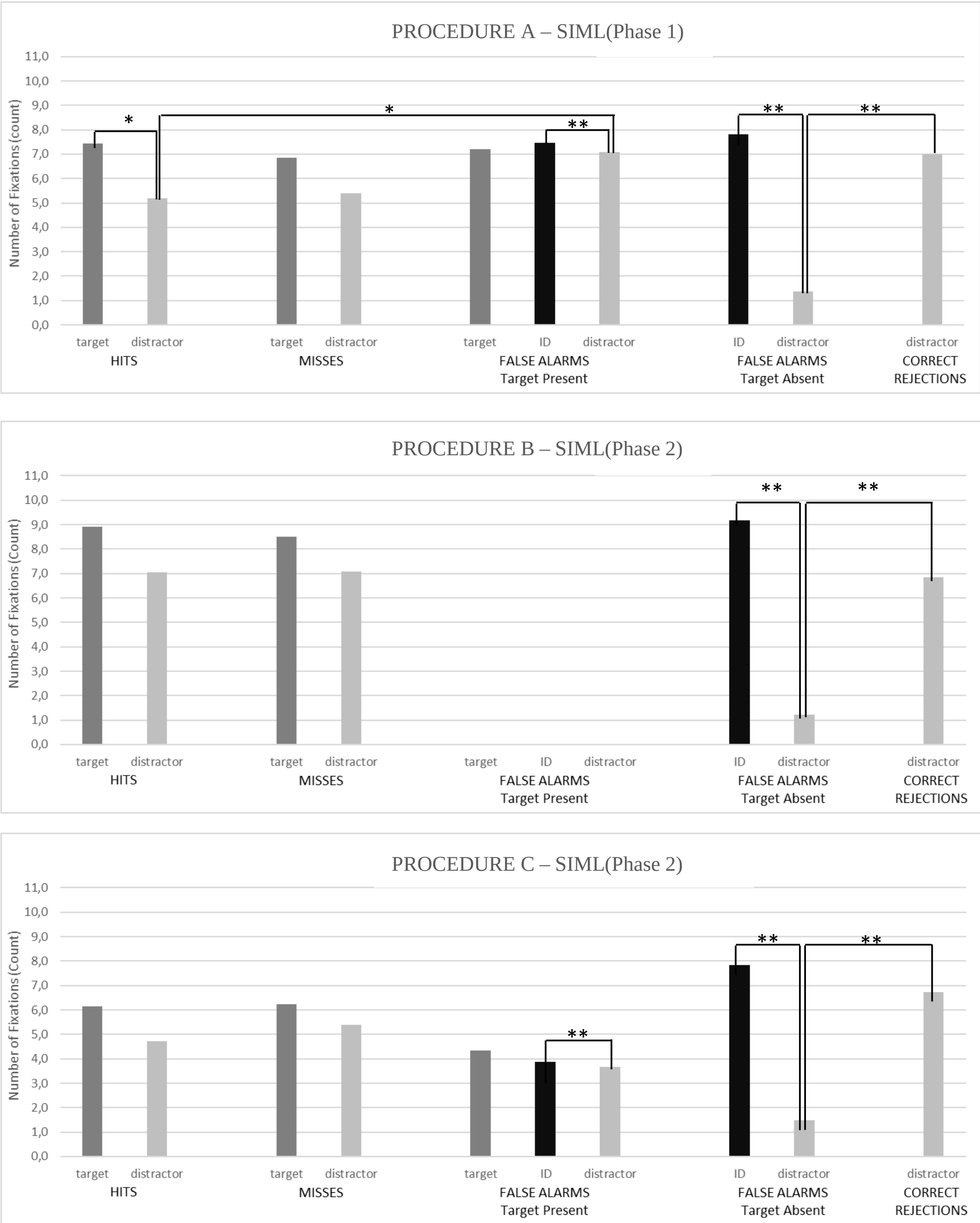


Figure 27. Overview of number of fixation results for all SIML Procedures

Facial features

Average dwell time on Internal and External Facial Features

For this analysis, the features were grouped into two categories: Internal Features (sum of dwell time on eyes, nose, and mouth) and External Features (sum of dwell time on hair, chin and ears).

Average dwell time on Internal and External Features – Phase 1 - Procedure A - SIML vs. Phase 1 - Procedure B - SEQL-StopR

The average dwell time on the internal and external features of the targets, distractors and identified individuals, in each response category, divided by procedure, is shown in Figure 28. Through visual data analysis of this Figure, we found that, for **Hits**, **Misses**, **False Alarms** on **Target Present** and **Correct Rejections**, the pattern of responses is very similar between procedures. However, a different pattern between procedures was observed for **False Alarms** on **Target Absent**. To explore this visual pattern a 2(type: identified vs distractor) x 2(features: Internal vs External) x 2(Procedure: A vs B - Phase 1) mixed ANOVA was carried out. The results revealed main effects of type ($F(1,36) = 22.821$, $p < .001$, $\eta_p^2 = .388$), features ($F(1,36) = 10.110$, $p < .01$, $\eta_p^2 = .219$), and procedure ($F(1,36) = 26.213$, $p < .001$, $\eta_p^2 = .421$). The following interactions were also observed: procedure*type ($F(1,36) = 389.507$, $p < .001$, $\eta_p^2 = .915$), type*features ($F(1,36) = 27.478$, $p < .001$, $\eta_p^2 = .433$), and a triple interaction type*features*procedure ($F(1,36) = 368.342$, $p < .001$, $\eta_p^2 = .911$). No interaction was obtained for features*procedure ($F(1,36) = .433$, $p = .515$, $\eta_p^2 = .012$). In fact, in procedure A, the average dwell time on the external characteristics of the distractor was higher ($t(18) = 7.94$, $p < .001$, Cohen's $d = 1.82$, CL effect size = .97) than on the internal characteristics. However, in procedure B, the same pattern was found for the identified individual ($t(18) = 8.55$, $p < .001$, Cohen's $d = 1.96$, CL effect size = .98).

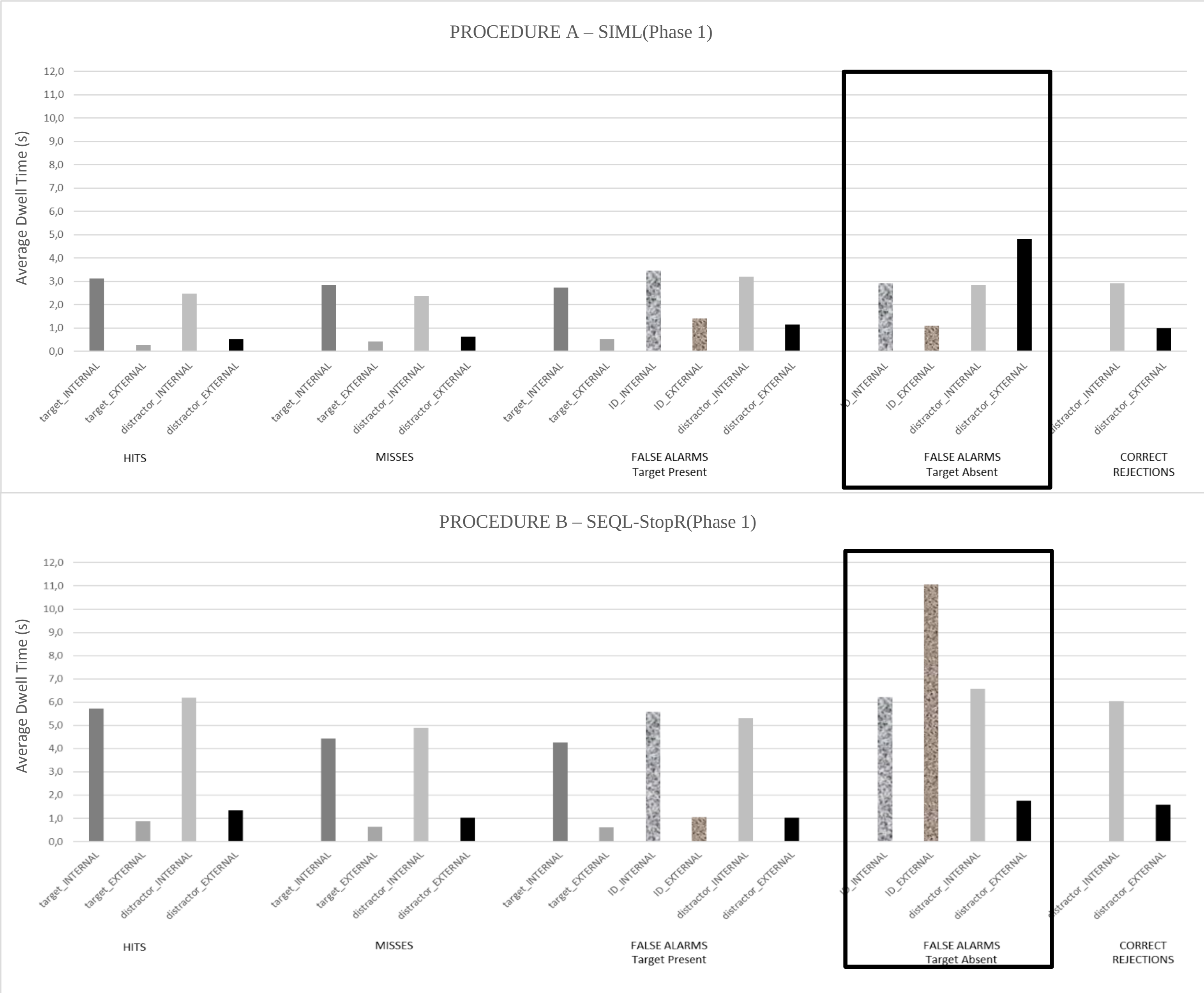


Figure 28. Average dwell time on Internal and External Face Features for Phase 1 - Procedure A - SIML vs. Phase 1 - Procedure B – SEQ-StopR.

Despite the similar patterns described above, other two 2(type: target vs. distractor) x 2(features: Internal vs External) x 2(Procedure: A vs B - Phase 1) ANOVA's were performed, one for **Hits** and the other for **Misses**. Both intended explore the small differences in dwell time on the features between targets and distractors. Concerning the **Hits**, results showed main effects of procedure, $F(1,35) = 19.57$, $p < .001$, $\eta_p^2 = .36$, and features, $F(1,35) = 286.08$, $p < .001$, $\eta_p^2 = .89$, but no effect of type was observed, $F(1,35) = 1.31$, $p = .260$, $\eta_p^2 = .04$. Additionally, features*procedure, $F(1,35) = 5.82$, $p = .021$, $\eta_p^2 = .14$, type*features, $F(1,35) = 7.91$, $p < .01$, $\eta_p^2 = .18$, and type*features*procedure, $F(1,35) = 6.06$, $p < .05$, $\eta_p^2 = .15$, interactions were observed.

Regarding the **Misses**, the results showed a main effect of features, $F(1,35) = 91.68$, $p < .001$, $\eta_p^2 = .83$, and procedure, $F(1,35) = 5.42$, $p < .05$, $\eta_p^2 = .22$, but no effect of type, $F(1,35) = .94$, $p = .343$, $\eta_p^2 = .05$, nor interactions.

Considering the opposite response patterns regarding **False Alarms** in the **Target Absent**, found between procedures and the slight differences observed for the **Hits** and the **Misses**, we decided to explore this data using all the face features (e.g., eyes, nose, ears) instead of groups (internal vs. external).

The average dwell time on each feature in the targets, distractors, and identified individuals, in the two procedures, A and B, Phase 1, when **Hits** occurred, are shown in Figure 29.

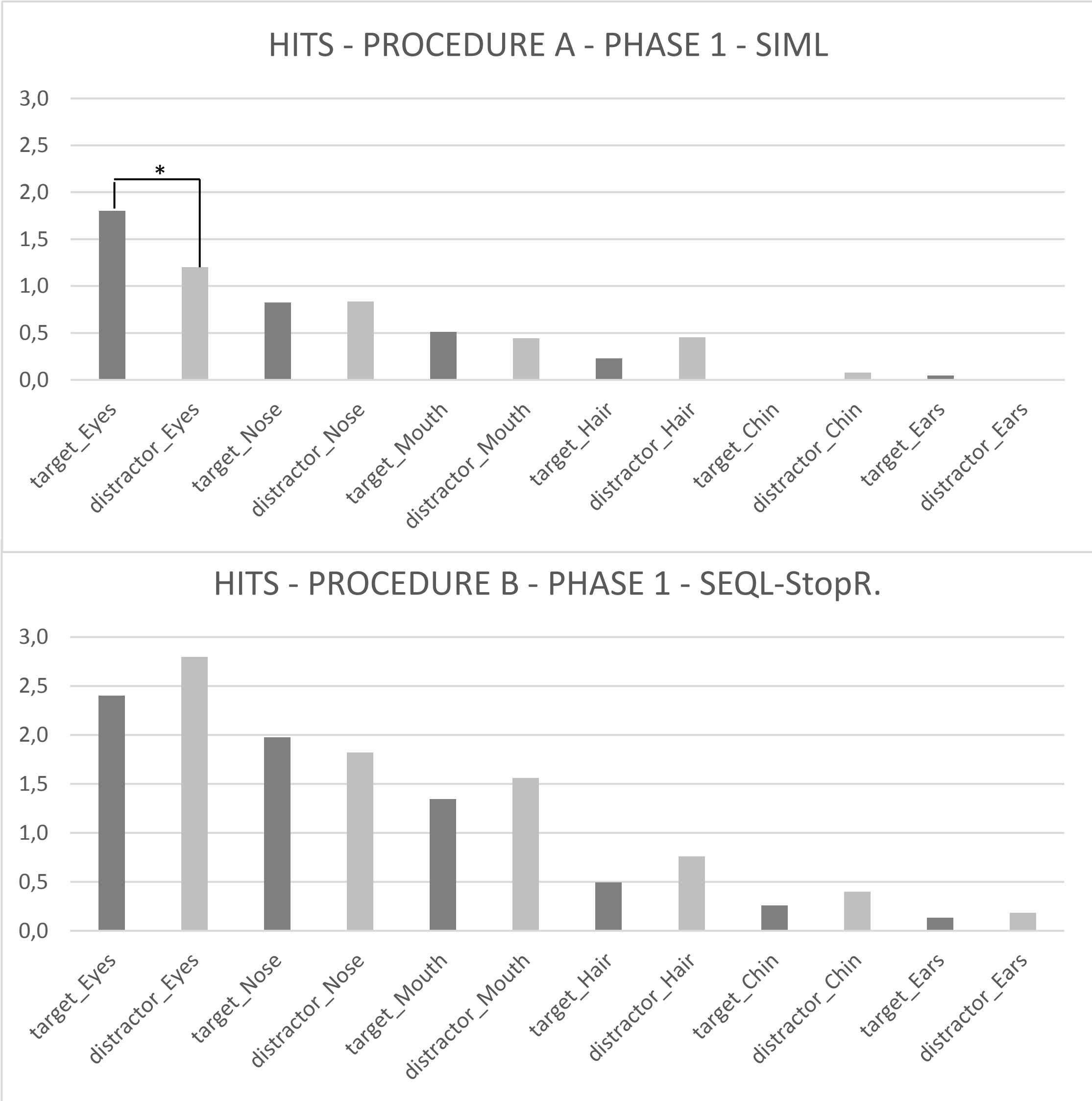


Figure 29. Average dwell time on each face feature in the targets, distractors, and identified non-target, in the two procedures, A and B, Phase 1, for Hits Responses.

In the category of **Hits**, a mixed ANOVA, 2 within (type: target vs. distractor) x 2 between (procedure: A vs. B), was performed for each feature. Regarding the **eyes**, results did not show any effect of type, $F(1,35) = .002$, $p = .962$, $\eta_p^2 = .00$, or procedure, $F(1,35) = 3.54$, $p = .07$, $\eta_p^2 = .09$, but an interaction was found $F(1,35) = 6.05$, $p = .019$, $\eta_p^2 = .15$. When we split by procedure, significant differences were only found for procedure A, $t(18) = 2.29$, $p < .05$, Cohen's $d = .52$, CL effect size = .70. For procedure B, when **Hits** took place, the average dwell time on the eyes of the targets was not significantly longer than in those of the distractors.

In relation to the **nose**, no effect of type, $F(1,35) = .97$, $p = .331$, $\eta_p^2 = .03$), or interaction, $F(1,35) = .19$, $p = .670$ $\eta_p^2 = .01$, were found. However, there was an effect of procedure, $F(1,35) = 19.83$, $p = .001$, $\eta_p^2 = .36$. Despite the absence of interaction, in a following split by procedure, in both procedures the average dwell time on the nose of the targets was significantly different from the average dwell time on the nose of the distractors.

In what regards the **mouth**, no effect of type, $F(1,35) = 1.04$, $p = .315$, $\eta_p^2 = .03$), or interaction, $F(1,35) = 2.20$, $p = .147$, $\eta_p^2 = .06$) were found, but an effect of procedure was obtained, $F(1,35) = 29.68$, $p < .001$, $\eta_p^2 = .46$. Analyzing separately, in both procedures the average dwell time on the mouth of the targets was significantly different from the average dwell time on the mouth of the distractors.

For **hair**, no effects of type, $F(1,35) = 2.40$, $p = .131$, $\eta_p^2 = .06$, procedure, $F(1,35) = .60$, $p = .445$, $\eta_p^2 = .02$), or interaction, $F(1,35) = .49$, $p = .490$, $\eta_p^2 = .01$, were found. Again, when we do a split by procedure, in all procedures the dwell time on the hair of the targets was significantly different from the dwell time on the hair of the distractors.

Regarding the **chin**, no effect of type, $F(1,35) = 2.60$, $p = .116$, $\eta_p^2 = .07$, or interaction, $F(1,35) = .09$, $p = .768$, $\eta_p^2 = .003$, were observed. However an effect of procedure was obtained, $F(1,35) = 4.31$, $p < .05$, $\eta_p^2 = .11$. When we separated the procedures (A and B), the dwell time on the chin of the targets was not significantly different from the dwell time on the chin of the distractors.

Concerning the **ears**, no effects of type, $F(1,35) = .01$, $p = .91$, $\eta_p^2 = .000$, procedure, $F(1,35) = 1.05$, $p = .313$, $\eta_p^2 = .03$, or interaction, $F(1,35) = 1.769$, $p = .192$, $\eta_p^2 = .048$, were found. The split by procedure showed that the dwell time on the ears of the targets was not significantly different from the dwell time on the ears of the distractors.

The total dwell time on each feature on the targets, distractors and identified non-target, in the procedures, A and B, 1st Phase, for the **Misses**, is in Figure 30.

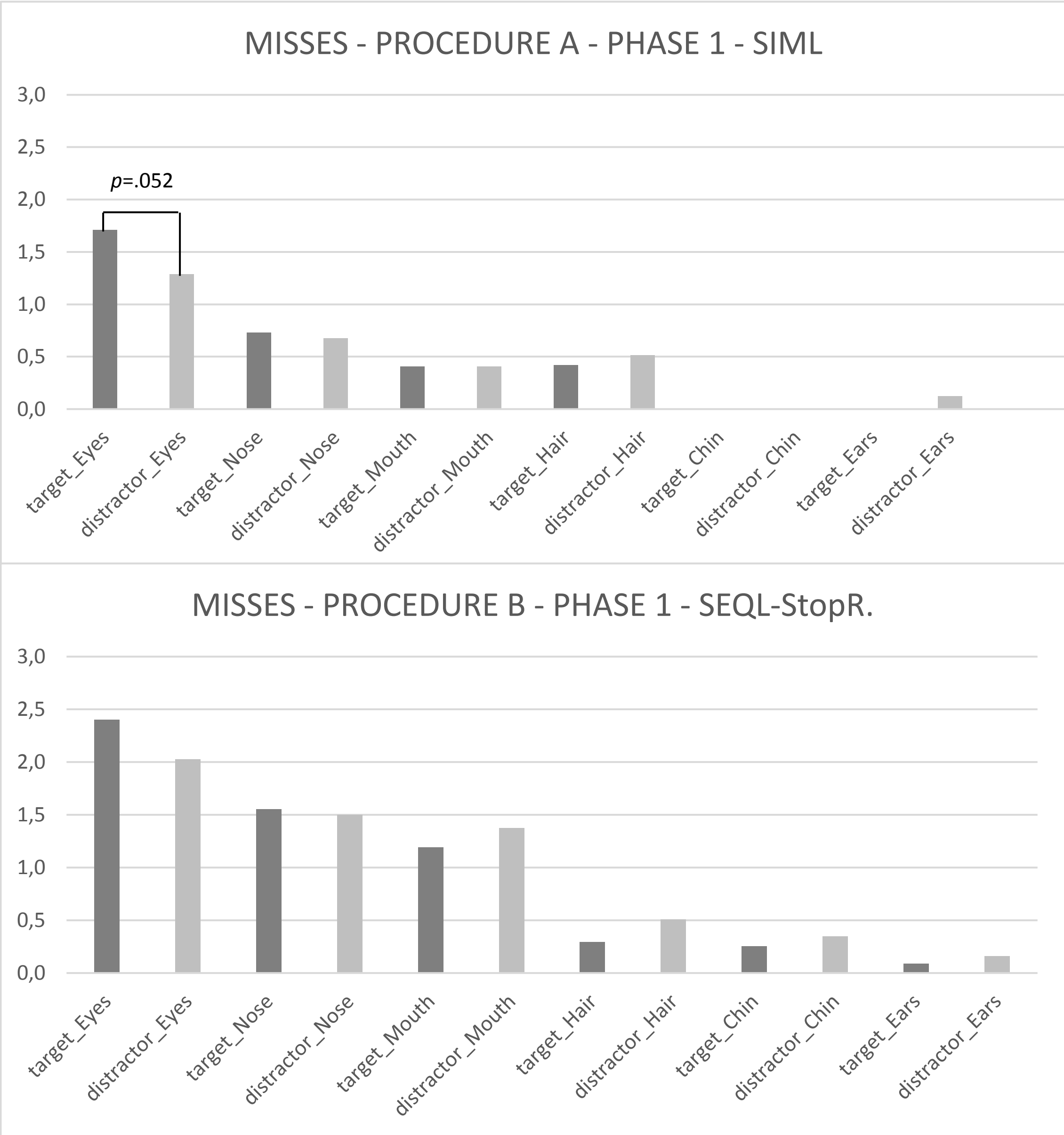


Figure 30. Average dwell time on each face feature in the targets, distractors, and identified non-target, in the two procedures, A and B, Phase 1, for Misses Responses.

Regarding to **Misses**, a mixed ANOVA, 2 within (type: target vs. distractor) x 2 between (procedure: A vs. B), was performed for each feature. About the **eyes**, there were no effects of type, $F(1,19) = .15$, $p = .701$, $\eta_p^2 = .08$, procedure, $F(1,19) = .07$, $p = .788$. $\eta_p^2 = .004$), or interaction, $F(1,19) = 3.53$, $p = .076$, $\eta_p^2 = .16$). In the split by procedure, only a marginal result ($p = .052$) was found for procedure A. Effectively, in **Misses** the average dwell time on the eyes of the targets was greater than the average dwell time on the eyes of the distractors, and this difference almost reached a significant result, $t(5) = 2.52$, $p = .052$, Cohen's $d = 1.03$, CL effect size = .80.

In relation to the **nose**, the results did not show effects of type, $F(1,19) = .001$, $p = .970$, $\eta_p^2 = .000$, or interaction, $F(1,19) = .07$, $p = .797$, $\eta_p^2 = .004$, but there was an effect of procedure, $F(1,19) = 12.84$, $p < .001$, $\eta_p^2 = .40$. When we do a split by procedure, for neither A or B the average dwell time on the nose of the targets was significantly different from the average dwell time on the nose of the distractors.

Concerning the **mouth**, no effects of type ($F(1,19) = .82$, $p = .376$, $\eta_p^2 = .04$) nor of interaction, $F(1,19) = .86$, $p = .365$, $\eta_p^2 = .04$, were found, but an effect of procedure, $F(1,19) = 17.60$, $p < .001$, $\eta_p^2 = .48$, was obtained. Analyzing separately, for none of the procedures the average dwell time on the mouth of the targets was significantly different from the average dwell time on the mouth of the distractors.

In regards to the **hair**, no effects of type, $F(1,19) = .81$, $p = .379$, $\eta_p^2 = .04$, procedure, $F(1,19) = .73$, $p = .405$, $\eta_p^2 = .001$, or of interaction, $F(1,19) = .02$, $p = .893$, $\eta_p^2 = .001$, were obtained. Again, when we do a split by procedure, for neither A or B the average dwell time on the hair of the targets was significantly different from the average dwell time on distractor's hair.

Concerning the **chin**, no effects of type, $F(1,19) = .51$, $p = .484$, $\eta_p^2 = .03$), procedure, $F(1,19) = 1.62$, $p = .218$, $\eta_p^2 = .08$, or of interaction, $F(1,19) = .51$, $p = .484$, $\eta_p^2 = .03$[23], were retrieved. When we separated the procedures (A and B) the average dwell time on the targets' chin was not significantly different from the average dwell time on the distractors' chin.

Referring to the **ears**, no effects of type, $F(1,19) = 2.04$, $p = .169$, $\eta_p^2 = .10$, procedure, $F(1,19) = .08$, $p = .775$, $\eta_p^2 = .004$, or of interaction, $F(1,19) = .24$, $p = .627$, $\eta_p^2 = .01$, were attained. The split by procedure showed that, both in A and B, the average dwell time on the ears of the targets was not significantly different from the average dwell time on the ears of the distractors.

[23] The same values recorded for the type. It is not an error, but the purest beauty of the magnificent random universe.

The average dwell time on each feature on the targets, distractors, and identified non-target, in the two procedures, A and B, 1st Phase, for the response category of **False Alarms** target absent, are in Figure 31.

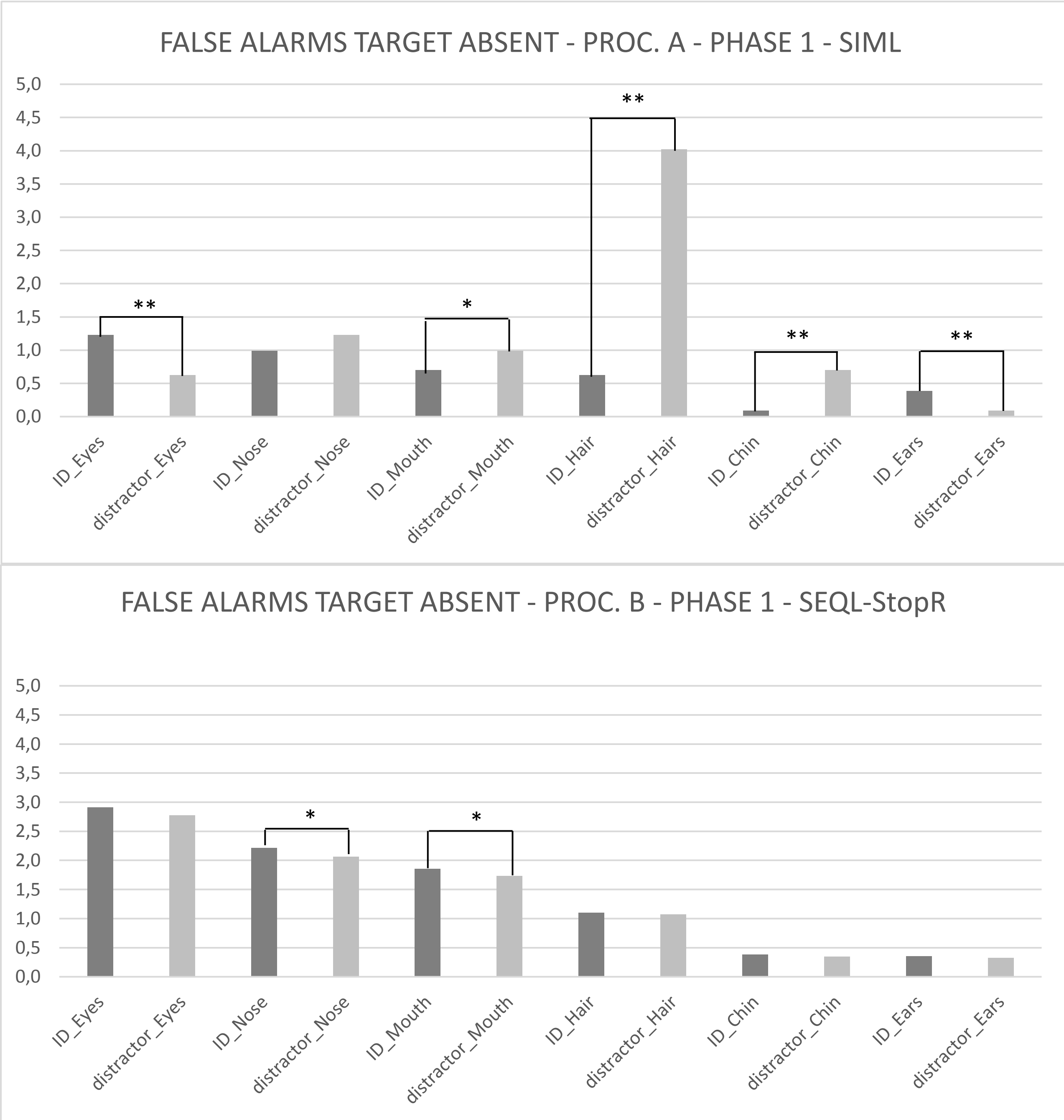


Figure 31. Average dwell time on each face feature on the targets, distractors and identified non-target, in the two procedures, A and B, Phase 1, for False Alarms Responses.

Relating to **False Alarms Target Absent**, a mixed ANOVA, 2 within (type: identified vs. distractor) x 2 between (Procedure: A vs. B), was performed for each feature. About the **eyes**, effects of type, $F(1,36) = 22.97$, $p < .001$, $\eta_p^2 = .39$, of procedure, $F(1,36) = 19.38$, $p < .001$, $\eta_p^2 = .35$, and interaction, $F(1,36) = 15.72$, $p < .001$, $\eta_p^2 = .30$, were retrieved. When we split by procedure, results show a significant difference in procedure A, but not in procedure B. The significant difference is related to the average dwell time on the eyes of the identified, that was greater than the average dwell time on the eyes of the distractor, $t(18) = 4.51$, $p < .001$, Cohen's $d = 1.06$, CL effect size = .86.

Regarding the **nose**, no effects of type, $F(1,36) = 1.83$, $p = .185$, $\eta_p^2 = .05$, procedure, $F(1,36) = 17.47$, $p < .001$, $\eta_p^2 = .33$, or interaction $F(1,36) = 5.35$, $p = .027$, $\eta_p^2 = .13$, were retrieved. Analyzing the procedures separately, results show a significant difference only in procedure B, specifically, the average dwell time on the nose of the identified was greater than the dwell time on the nose of the distractor, $t(18) = 2.17$, $p < .05$, Cohen's $d = .51$, CL effect size = .69.

In relation to the **mouth**, effects of type, $F(1,36) = 5.12$, $p = .030$, $\eta_p^2 = .13$, procedure, $F(1,36) = 31.65$, $p < .05$, $\eta_p^2 = .001$, and interaction, $F(1,29) = 10.29$, $p < .01$, $\eta_p^2 = .22$ were obtained. In a split by procedure, significant differences were found in both procedures in average dwell time on the mouth of the distractors and of the identified. In procedure A, the average dwell time on the mouth of the distractor was greater than the average dwell time on the mouth of the identified, $t(18) = 2.80$, $p < .05$, Cohen's $d = .66$, CL effect size = .75. In procedure B, results revealed that the average dwell time on the mouth of the identified was greater than the average dwell time on the mouth of the distractor, $t(18) = 2.16$, $p < .05$, Cohen's $d = 0.51$, CL effect size = .69.

Respecting **hair,** effects of type, $F(1,36) = 220.89$, $p < .001$, $\eta_p^2 = .86$, procedure, $F(1,36) = 8.11$, $p = .01$¸ $\eta_p^2 = .18$, and an interaction, $F(1,36) = 225.72$, $p < .001$, $\eta_p^2 = .86$, were found. Analyzing separately, significant differences were found only in procedure A, where the average dwell time on the distractors' hair was greater than the average dwell time on the hair of the identified, $t(18) = 15.15$, $p < .001$, Cohen's $d = 3.57$, CL effect size = .99.

Regarding the **chin**, an effect of type, $F(1,36) = 32.63$, $p < .001$, $\eta_p^2 = .48$ and an interaction, $F(1,36) = 35.99$, $p < .001$ $\eta_p^2 = .50$, were retrieved, but no effect of procedure, $F(1,36) = .47$, $p = .499$, $\eta_p^2 = .01$) was observed. When we separated the procedures, procedure A showed significant differences, since the average dwell time on the chin of the distractor was greater than the average dwell time on the chin of the identified, $t(18) = 5.88$, $p < .001$, Cohen's $d = 1.38$, CL effect size = .917.

In relation to the **ears**, the results revealed an effect of type, $F(1,36) = 11.13$, $p < .01$, $\eta_p^2 = .24$, and an interaction, $F(1,36) = 9.64$, $p < .001$, $\eta_p^2 = .21$, but no effect of procedure, $F(1,36) = .06$, $p = .803$, $\eta_p^2 = .002$. The split by procedure showed significant differences only in procedure A. Indeed, the average dwell time on the distractors' ears was greater than the dwell time on the ears of the identified, $t(18) = 3.23$, $p < .001$, Cohen's $d = .76$, CL effect size = .77.

Average dwell time on Internal and External Features – Phase 1, Procedure A – SIML vs. Phase 2, Procedure B – SIML after SEQL-StopR

The average dwell time on the internal and external characteristics of the targets, distractors and identified non-targets, in each response category, divided by procedure, is shown in Figure 32.

Through visual data analysis of Figure 32, we find that for the all the categories, the pattern of responses match between procedures. A 2(Type: target vs. distractor) x 2(Features: Internal vs. External) x 2(Procedure: A - Phase 1 vs. B - Phase 2) ANOVA was performed for **Misses**, to explore the slight differences of dwell time in features between targets and distractors.

The results showed a main effect of feature, $F(1,15) = 130.28$, $p < .001$, $\eta_p^2 = .90$, and a statistical trend of procedure, $F(1,15) = 4.31$, $p = .056$, $\eta_p^2 = .22$, but no effect of type, $F(1,15) = 1.33$, $p = .266$, $\eta_p^2 = .08$. More, all the interactions were observed (*ps* > .135). Importantly, the dwell time in external features of distractor in Procedure B SIML is significantly greater than in Procedure A ($t(15) = -2.57$, $p = .021$, Cohen's $d = 1.24$, CL effect size = .81).

Due to the lack of results in the response category **False Alarms – Target Present**, the features' detailed analyses are not presented.

Visual data analysis of Figure 32 also shows that for **False Alarms** in **Target Absent**, the dwell time in external features of the distractor in Procedure B SIML is greater than in SIML of Procedure A. The *t*-test shows that this difference achieved a statistical trend ($t(29) = -1.94$, $p = .069$, Cohen's $d = .95$, CL effect size = .75).

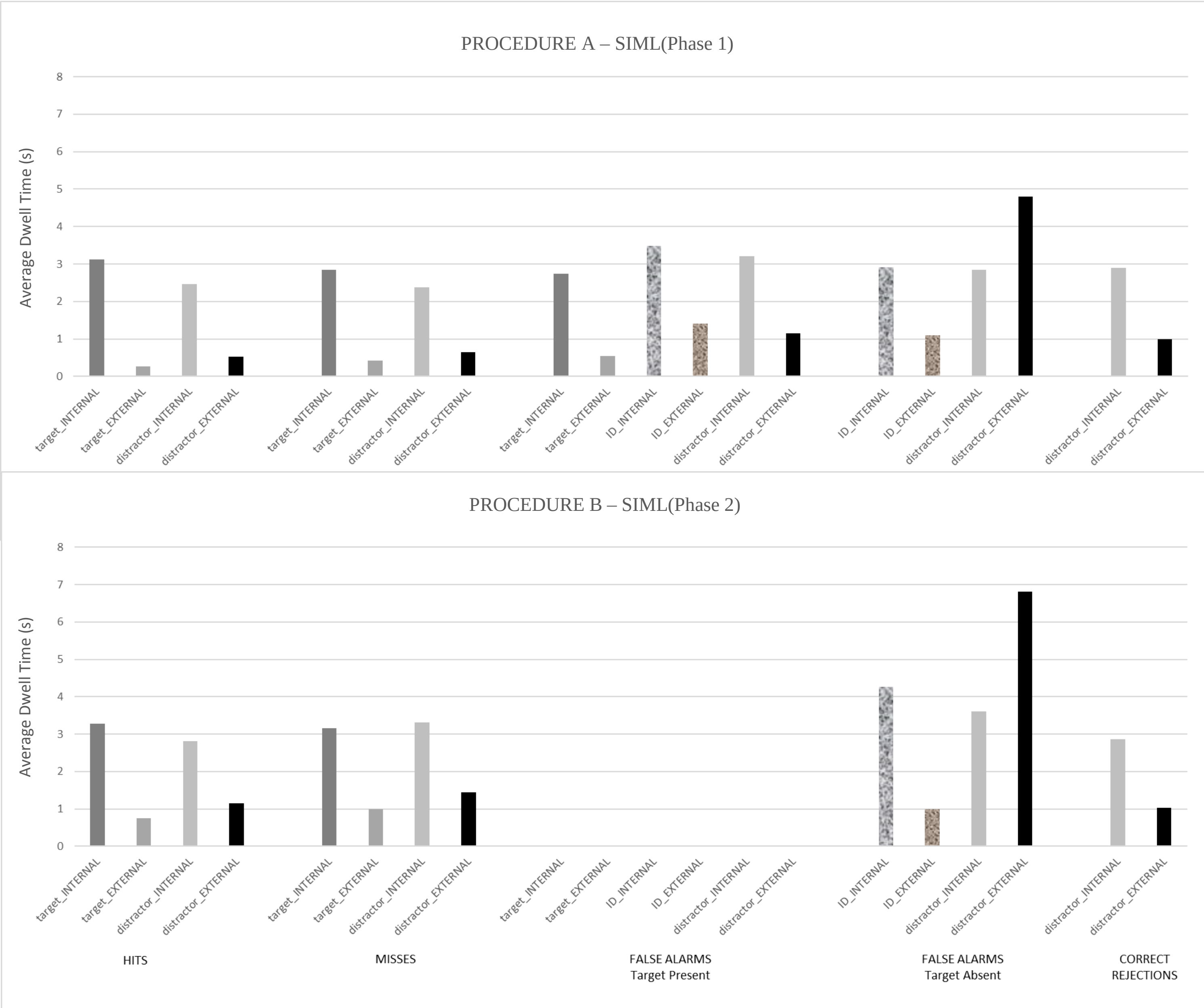
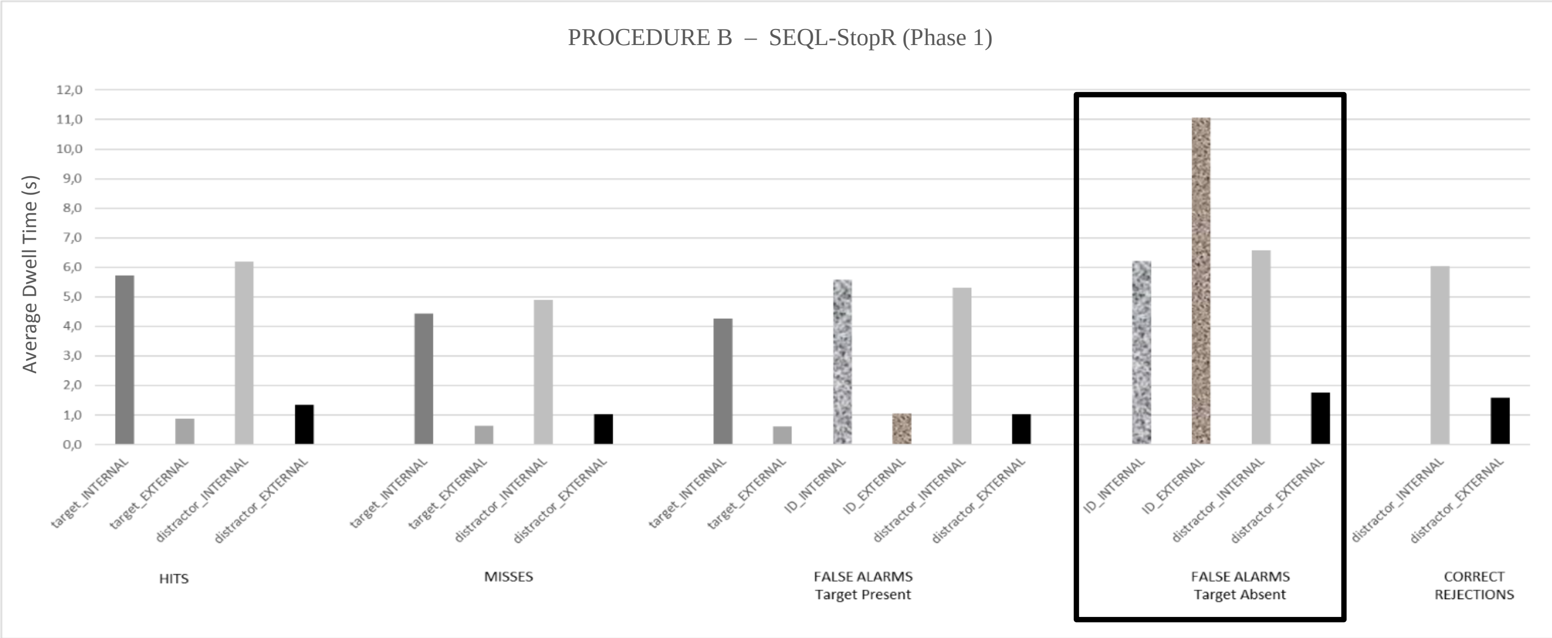


Figure 32. Average dwell time on Internal and External Face Features for Phase 1, Procedure A – SIML vs. Phase 2, Procedure B - SIML.

Average dwell time on Internal and External Features – Phase 1 - Procedure B - SEQL-StopR vs. Phase 2 - Procedure B - SIML after SEQL-StopR

The average dwell time on the internal and external features of the targets, distractors and identified non-targets, in each response category, divided by procedure, is shown in Figure 33. Through the visual data analysis of Figure 33, we found that, for the **Misses**, **False Alarms** on **Target Present** and **Correct Rejections**, the pattern of responses fits almost entirely between procedures. However, a difference between procedures pattern was observed for **Hits**, and specially for the **False Alarms** on **Target Absent**. To explore the **Hits** pattern, a 2(Type: Target vs Distractor) x 2(Features: Internal vs External) x 2(Phase 1: SEQL vs Phase 2: SIML) ANOVA was produced. The results revealed an effect of features, $F(1,17) = 117.68$, $p < .001$, $\eta_p^2 = .87$, phase, $F(1,17) = 5.88$, $p = .027$, $\eta_p^2 = .26$, and an interaction features*phase $F(1,17) = 24.92$, $p < .001$, $\eta_p^2 = .59$. However no effects of type, $F(1,17) = 1.29$, $p = .277$, $\eta_p^2 = .07$, or other interactions were registered (*ps* > .143).

Regarding the **False Alarms** in **Target Absent** a 2(Type: ID vs Distractor) x 2(Features: Internal vs External) x 2(Phase 1: SEQL vs Phase 2: SIML) ANOVA was produced. The results revealed main effects for type, $F(1,11) = 5.55$, $p < .05$, $\eta_p^2 = .34$, features, $F(1,11) = 5.30$, $p < .05$, $\eta_p^2 = .33$, phase, $F(1,11) = 7.75$, $p < .05$, $\eta_p^2 = .34$, and type*phase, $F(1,11) = 102.07$, $p < .001$, $\eta_p^2 = .90$, and a type*features*phase, $F(1,11) = 120.22$, $p < .001$, $\eta_p^2 = .92$, interactions. No other interactions were found (*ps* > .281). Given these results, we decided to explore this data using all the face features (e.g., eyes, nose, ears) instead of groups (internal vs. external)



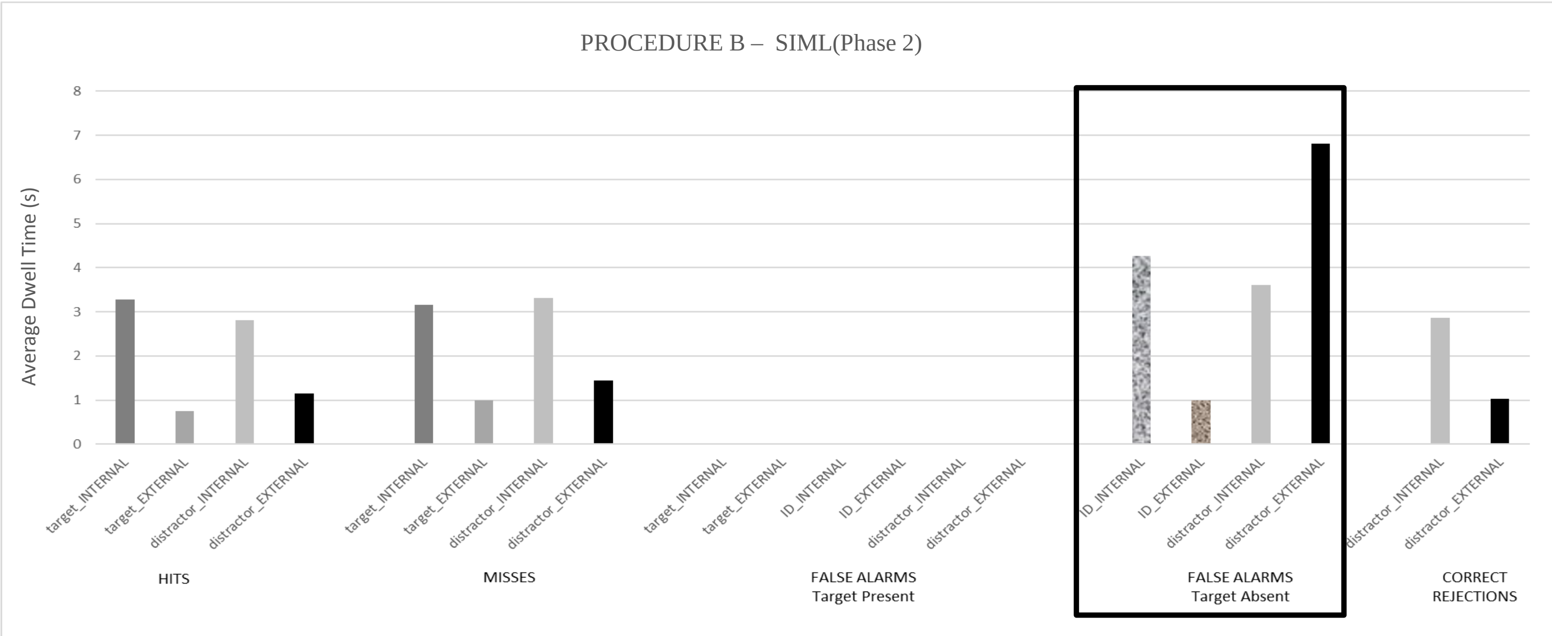


Figure 33. Average dwell time on Internal and External Face Features for Phase 1 - Procedure B – SEQL-StopR vs. Phase 2 - Procedure B - SIML.

The total dwell time on each feature on the targets, distractors and identified non-target, in the two phases of procedures B, for the response category ‘**Hits**’, is in Figure 34.

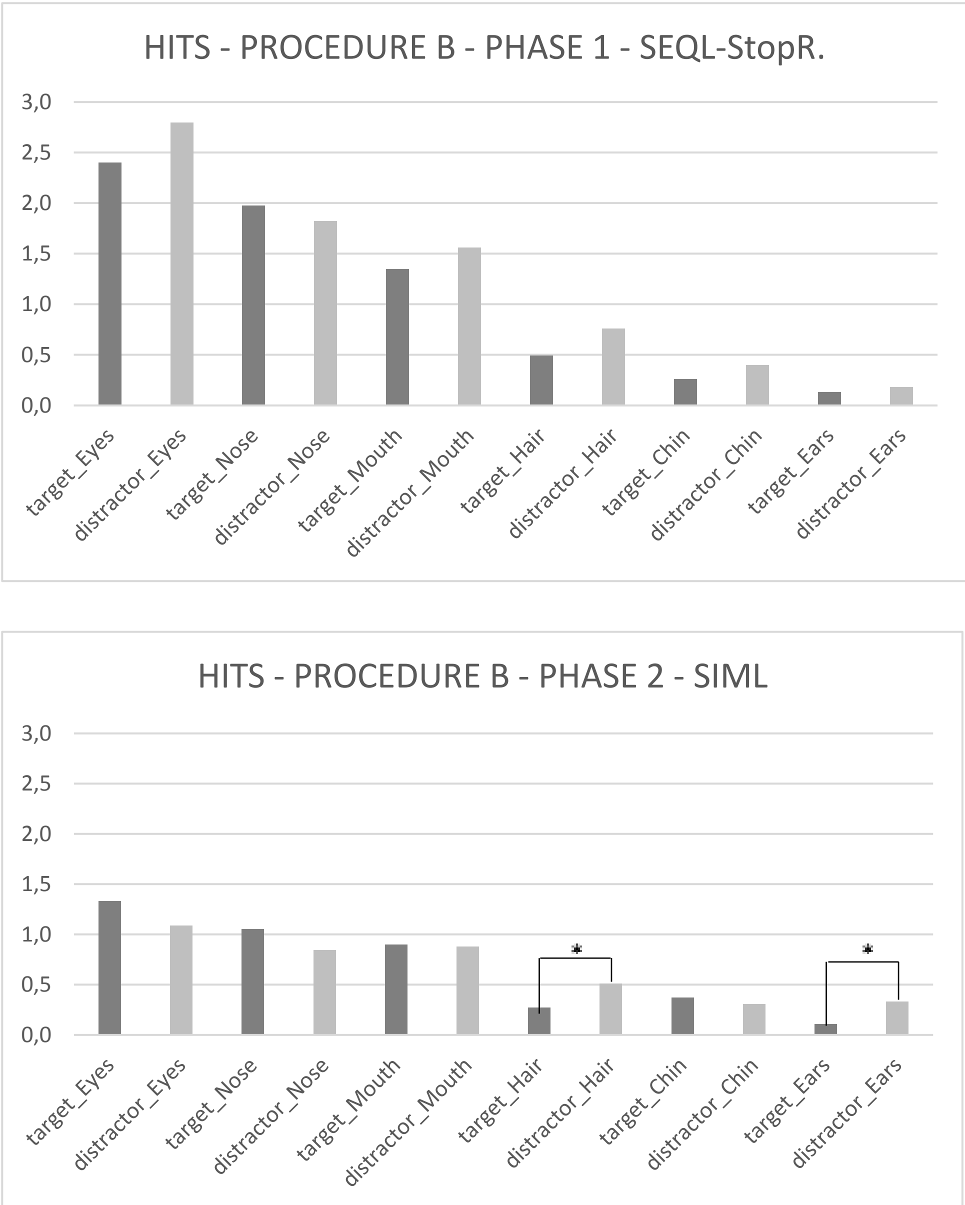


Figure 34. Average dwell time on each face feature in the Identified suspect and distractors in the two Phases of procedures B, for Hits Responses.

With reference to the **Hits,** a repeated measures ANOVA, 2 within (type - Target vs Distractor) x 2 within (phase/lineup - SEQL-StopR B vs. SIML B), was performed for each feature. Relating to the **eyes**, the results expressed an effect of phase, $F(1,17) = 10.81$, $p < .001$, $\eta_p^2 = .39$, but no effect of type, $F(1,17) = .79$, $p = .386$, $\eta_p^2 = .04$, nor of interaction, $F(1,17) = 2.54$, $p = .129$, $\eta_p^2 = .13$.

Concerning to the **nose**, an effect of phase, $F(1,17) = 16.51$, $p < .001$, $\eta_p^2 = .49$, was obtained. However no effect of type, $F(1,17) = .24$, $p = .632$, $\eta_p^2 = .01$, nor of interaction, $F(1, 17) = .005$, $p = .942$, $\eta_p^2 = .000$, were reached.

A similar pattern of results was found for the **mouth**. The ANOVA showed an effect of phase, $F(1,17) = 10.94$, $p < .01$, $\eta_p^2 = .39$, a statistical trend of type, $F(1,17) = 3.95$, $p = .063$, $\eta_p^2 = .11$, and absence of interaction, $F(1,17) = .170$, $p = .685$, $\eta_p^2 = .01$.

In what regards the **hair**, no effects of phase, $F(1,17) = .09$, $p = .773$, $\eta_p^2 = .01$, type, $F(1,17) = 2.32$, $p = .146$, $\eta_p^2 = .12$, or interaction, $F(1,17) = .99$, $p = .335$, $\eta_p^2 = .06$, were obtained. Splitting by procedure, results showed a significant difference only in Phase 2 of procedure B. Indeed, the average dwell time on the hair of the distractor was greater than the average dwell time on the hair of the target, $t(19) = 2.38$, $p < .05$, Cohen's $d = .53$, CL effect size = .70.

For the **chin**, the results did not reveal any effects of type, $F(1,17) = .684$, p = .430, $\eta_p^2 = .039$, Phase, $F(1,17) = .45$, $p = .512$, $\eta_p^2 = .06$, nor of interaction, $F(1,17) = .58$, $p = .46$, $\eta_p^2 = .03$.

Relatively to the **ears**, results showed an effect of type, $F(1,17) = 9.70$, $p < .01$ $\eta_p^2 = .36$, but no effects of phase, $F(1,17) = 1.39$, $p = .255$, $\eta_p^2 = .08$, or interaction, $F(1,17) = 2.50$, $p = .133$, $\eta_p^2 = .13$. Splitting by procedure, results show a significant difference only in Phase 2 of procedure B. The average dwell time on the ears of the distractor was greater than the average dwell time on the ears of the target, $t(19) = 2.91$, $p < .05$, Cohen's $d = .65$, CL effect size = .74.

The total dwell time on each feature on the distractors and identified non-target, in the two phases of procedure B, for the response category of '**False Alarms** in target Absent, is shown in Figure 35.

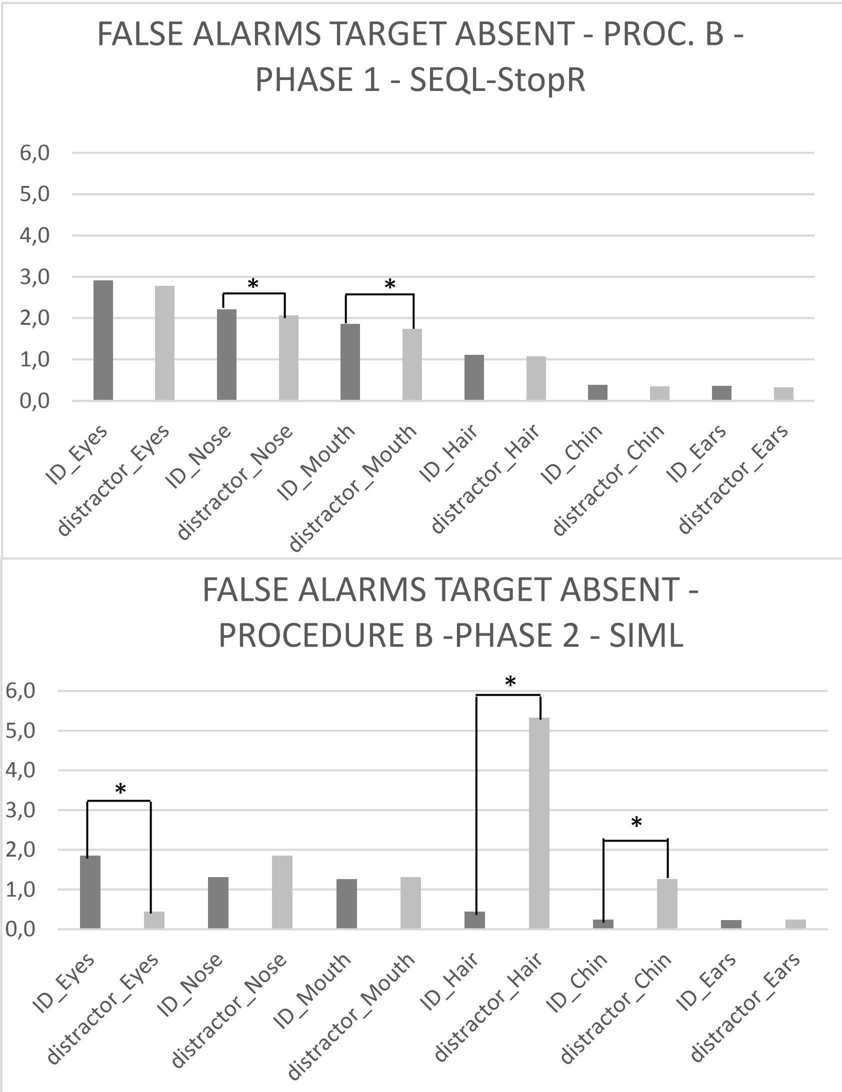


Figure 35. Average dwell time on each face feature in the Identified suspect and distractors in the two Phases of procedure B for 'False Alarms –Target Absent' Responses.

In relation to '**False Alarms - Target Absent'**, a repeated measures ANOVA, 2 within (type - Target vs Distractor) x 2 within (phase/lineup - SEQL-StopR B vs. SIML B), was performed for each feature. Concerning the **eyes**, the results reveal an effect of type $F(1,11) = 12.48$, $p < .001$, $\eta_p^2 = .53$, phase, $F(1,11) = 11.87$, $p < .01$, $\eta_p^2 = .52$, and interaction, $F(1, 11) = 12.42$, $p < .001$, $\eta_p^2 = .53$.

About the **nose**, the opposite pattern was observed. No effects of type, $F(1,11) = .31$, $p = .587$, $\eta_p^2 = .03$, phase, $F(1,11) = 2.48$, $p = .144$, $\eta_p^2 = .18$, or interaction, $F(1,11) = 1.53$, $p = .242$, $\eta_p^2 = .12$, were retrieved.

Regarding the **mouth**, an effect of phase was found, $F(1,11) = 5.32$, $p < .05$, $\eta_p^2 = .33$. Notwithstanding no effects of type, $F(1,11) = .179$, $p = .680$, $\eta_p^2 = .02$, or interaction, $F(1,11) = 1.28$, $p = .28$, $\eta_p^2 = .10$.

For the **hair**, an effect of type, $F(1,11) = 88.64$, $p < .001$, $\eta_p^2 = .89$, an interaction $F(1,11) = 89.40$, $p < .001$, $\eta_p^2 = .89$, and a statistical trend of phase, $F(1,11) = 4.15$, $p = .07$, $\eta_p^2 = .27$, were obtained.

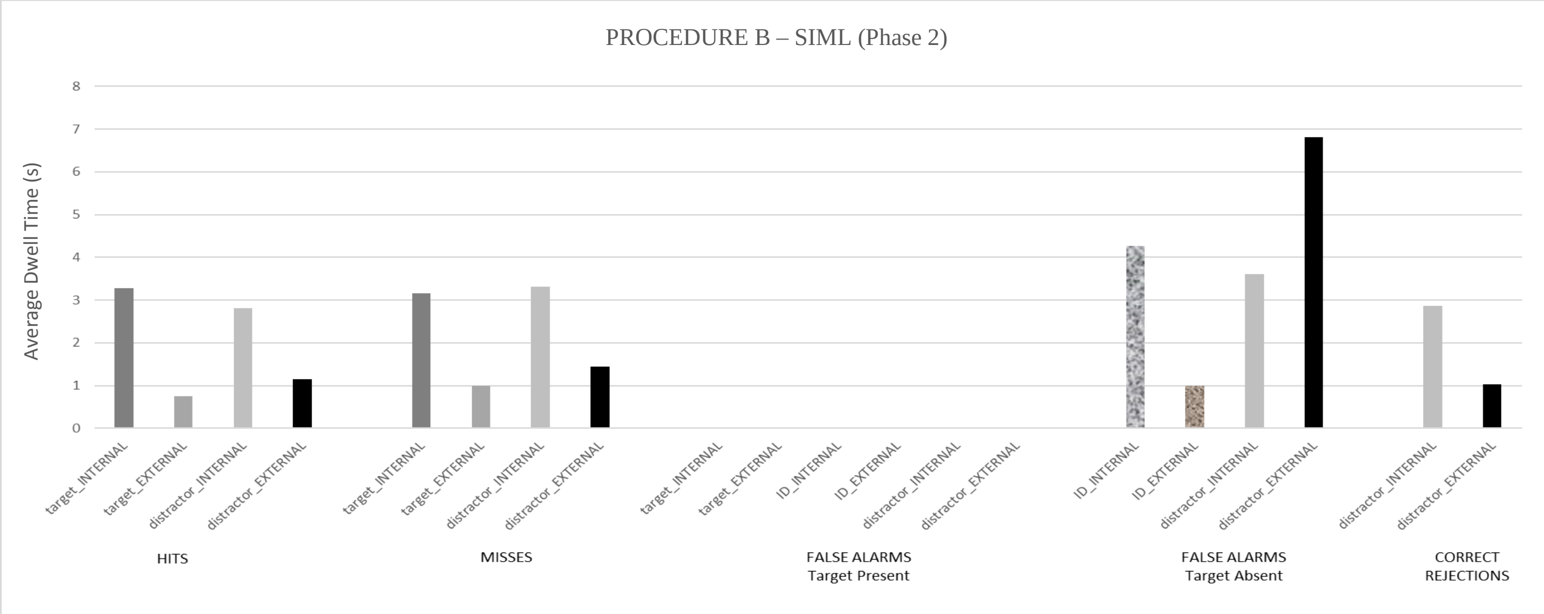
Respecting the **chin**, the results revealed an effect of type, $F(1,11) = 27.55$, $p < .001$, $\eta_p^2 = .72$, phase, $F(1,11) = 20.46$, $p < .001$, $\eta_p^2 = .65$), and interaction, $F(1,11) = 28.97$, $p < .001$, $\eta_p^2 = .73$.

In relation to the **ears**, no effects of type, $F(1,11) = .00$, $p = .999$, $\eta_p^2 = .00$, phase, $F(1,11) = .08$, $p = .778$, $\eta_p^2 = .008$, or interaction, $F(1,11) = .01$, $p = .940$, $\eta_p^2 = .001$, were reached.

Average dwell time on Internal and External Features – Phase 2 - Procedure B - SIML vs. Phase 2 - Procedure C - SIML

The average dwell time on the internal and external features of the targets, distractors and identified non-targets, in each response category, divided by procedure, is shown in Figure 36. Through the visual data analysis of Figure 36, we found that, for the **Hits**, **Misses**, **False Alarms** on **Target Present** and **Absent** and **Correct Rejections**, the pattern of responses fit entirely between procedures and no inferential analyzes were performed.

Due to the proximity of the face features and the possible error in saccadic movements the analysis for fixations were not performed.



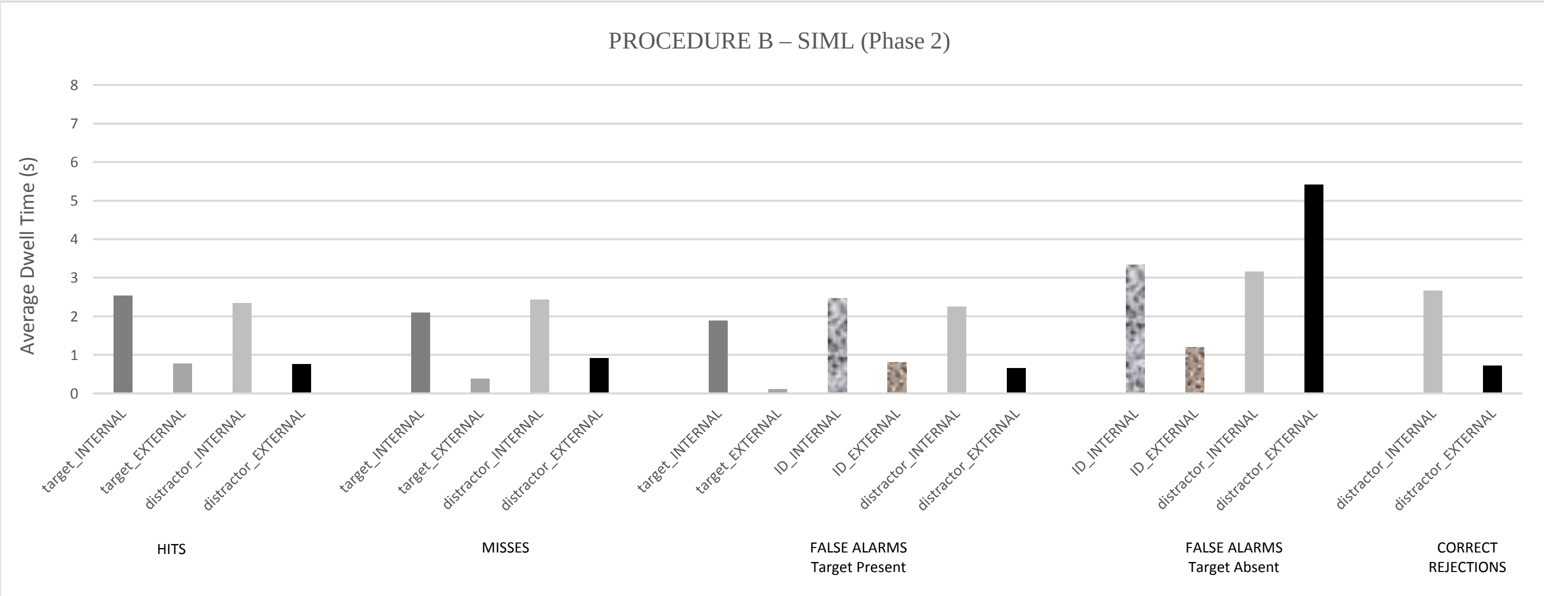


Figure 36. Average dwell time on Internal and External Face Features for Phase 2 - Procedure B – SIML vs. Phase 2 - Procedure C – SIML.

DISCUSSION

The errors associated with eyewitness testimony have received more than 100 years of scientific attention (e.g., Münsterberg, 1908). Although many studies had a major impact on the judicial system, the study of Lindsay and Wells in 1985 brought a new perspective to procedures used in lineups. Having a strong experimental theory supporting it, SEQL was asserted by allowing an absolute judgment, which reduces the number of **False Alarms**/Positives, by promoting the comparison of each suspect with the real memory of the criminal (Steblay et al., 2011; Pozzulo et al., 2009). However, if it is true that SEQL provides fewer **False Alarms**/Positives than the SIML, it also provides a lower rate of culprit identification (Mecklenburg, 2006; Mansour & Flowe, 2010; Levi, 2016). Considering this, Malpass (2006) and Wells (2014) presented questions to which the answers they considered to be fundamental. Malpass (2006) asked should we increase the probability of identifying the perpetrator at the risk of increasing the possibility of mistaken identification? Alternatively, should mistaken identifications be avoided at all costs, enabling a criminal to escape sentence? In the same line Wells (2014, p. 14) inquired "Should we adopt a new procedure (SIML) that increases the chances that the guilty might be identified but also increases the chances of mistaken identification?".

The main objective of this study was to verify if the fusion of the two procedures (SEQL and SIML) could somehow bring about the 'best of two worlds'. That is, if in a first moment, the SEQL could promote a low number of **False Alarms**/Positives and, in a second phase/moment, maintaining the effect of absolute judgment, the SIML could facilitate increased identifications of the culprit (identification of the criminal in the lineup), and a decrease in **Misses** (situation where the eyewitness rejects lineup when the criminal was there - leaving the criminal in the society). Using eyetracking measures we wanted to explore these carry-over effects and respond to Garry Wells' challenge (2014), exploring markers of correct identifications. Due to the lack of literature and to the absence of robust results in previous studies, we also tried to replicate some of the results found in previous eyetracking studies.

For this, three procedures were developed. Procedure A consisting of a classic SIML, Procedure B consisting of a SEQL-StopR, in the first phase, followed by a SIML, in the second phase, and Procedure C consisting of SEQL-Pass (no formal decision was required, merely visualizing each face for 5 seconds), in a first phase, followed by a SIML in the second phase.

Behavioural Measures

Behavioural results showed that, in the first phase of procedures A and B (SIML vs. SEQL-StopR), participants who were facing a SEQL-StopR gave significantly more **Misses**, but significantly fewer **False Alarms**/Positives in **target absent** conditions than those who visualized the SIML. This result is in line with the literature (Mecklenburg, 2006; Levi, 2016), and was one of the leading hypotheses of the present study. Considering the literature in eyewitness testimony, it was also expected that the SIML promoted a greater number of **Hits**. In fact, this result was achieved, but even though these differences are around ten percentage points, no statistical significance was achieved.

Comparing the two phases of procedure B (SEQL-StopR vs. SIML after SEQL - within-subjects comparison) the participants in the second phase gave significantly more **Hits** in the target, fewer **False Alarms** in **target present** condition, and an equal number of **False Alarms** in the **target absent** condition. This result is mostly in unison with what was expected, since performing the SIML after the SEQL-StopR allowed for a greater number of **Hits** in the target, without compromising the gains obtained by the SEQL in the first phase, namely at the level of the **False Alarms** in the **target absent**. In fact it was not expected that in the second phase (SIML) there would be fewer **False Alarms** in the **target present**, since the SIML, by providing a relative judgment and promoting a comparative judgement, is more susceptible to **False Alarms** (Lindsay et al., 1991; Steblay et al., 2011; Wells et al., 2009; Wells et al., 2006). It was, however, expected that, by performing a SEQL in the first phase, the number of **false alarm errors** in either **target absent** and **target present** conditions would be maintained. This unexpected result for the **target present** can be explained by the fact that the present study was carried out with SEQL-StopR. Effectively, when the participant is facing a SEQL-StopR, if he identifies someone, he cannot see any more suspects and cannot go back. This may have meant that the participants who identified an innocent in the first phase (**target present**) might have done so based on a weak memory trace, with the aggravating fact of not having seen all the suspects including the target (precisely in the **target present** case). Later, when they were shown the SIML, and the target was present, the participant was allowed the visualization of all stimuli, target and distractors, and to make the relative comparisons, allowing an identification of the criminal, since the SIML promotes the correct identification of the target when he is present (Lindsay, Mansour, Beaudry, Leach, & Bertrand, 2009a; Mickes, Flowe, & Wixted, 2012). In consonance with this, in one of the reference works in eyewitness research, Steblay, Dysart, Fulero, & Lindsay, (2001, p. 469), which states that "SIML generate a higher correct identification rate by leading some witnesses who have a weak memory trace to choose anyway. Because the target is the best match to their memory (on average), these guesses are somewhat better than chance, and a higher rate of target choices is obtained". Thus, when the target was not present, this process was not possible and the previous SEQL, despite the relative judgment and the increase of the liberal criterion characteristic of SIML (Gronlund et al., 2012), made it possible for there not to be a significant increase in the number of **False Alarms** on the **target absent** condition in the second phase.

Comparing the SIML of procedure A with the SIML (2^{nd} phase) of procedure B, when individuals viewed the second phase of procedure B, they gave significantly more **Hits** and gave significantly fewer **False Alarms**/Positives in **target present**. These individuals who observed SIML after the SEQL (procedure B) tended to have less **False Alarms** on **target absent**, and **Correct Rejections**, but also had more **Misses**. With the exception of the **Misses**, the results seem to corroborate what was expected from the literature. Indeed, the individuals who performed a SIML after the SEQL appeared to opt for a mixed strategy, balanced between conservative and liberal, allowing for the gains of the absolute judgment to be maintained, with a posterior judgment that allows greater liberality and comparison, facilitating identification when the memory trace is weak (Flowe, Mehta, & Ebbesen, 2011; Meissner, Tredoux, Parker, & MacLin, 2005). The fact that the **Misses** do not decrease significantly in the visualization of an SIML after a SEQL (procedure B) can be explained by the argument that the previous SEQL has promoted a more conservative choice in the SIML, not allowing for the increase of identification answers (that would reduce the number of **Misses**), but that would possibly also increase the number of **False Alarms** (Flowe & Ebbesen, 2007; Meissner et al., 2005). In fact, the individuals who made the **Misses** in the first phase of procedure B had already visualized all the faces with exhaustive exploration, having rejected the lineup. Hence, when they visualized the lineup in the next phase (SIML), the conservative strategy used in the first phase contaminates the relative judgment and the liberal strategy of the second.

Concerning the comparison of the second phases of procedures B and C, the results show that individuals who visualize the SIML of procedure B hit significantly more on the target (HIT), and give significantly fewer **False Alarms**/Positives in the **target present**. Interestingly, although there is a prior SEQL in procedure C, unlike in procedure B in which the participant has to make a decision and can explore the faces unlimitedly, the SEQL-Pass has a visual exploration limited to 5s for each face and does not require formal decision making. The differences observed between the number of **False Alarms** on **target present** in the SIML of procedures B and C may be due to the fact that, although they did not make a formal judgment, the participants who performed the SEQL-Pass knew that they would see all the suspects and that all these suspects would appear later in the SIML. Thus, in the SEQL-Pass the participants may have performed, albeit involuntarily, an absolute judgment (comparing with the memory that they had of the criminal), but also a relative judgment, comparing the present suspect with the suspects who have already been seen in the lineup. Therefore, since the participants performing procedure C had already been exposed to the criminal and all suspects in the first phase, the impact of the SIML of phase 2 is not as positive as in procedure B.

The behavioural results were, therefore, supportive of our hypothesis that a mixed procedure, consisting of a SIML after a SEQL-StopR, can combine the best characteristics of the two lineups, allowing for better performance. In fact, there seems to be a positive contamination of the absolute strategy typical of the SEQL to the relative strategy of the second phase SIML. Based on the literature of dual processes, which postulates the existence of two distinct processes that are related to independent cognitive operations, empirical studies in eyewitness testimony show that the process of recollection (which reflects the retrieval of qualitative information about a specific study episode -Yonelinas, Aly, Wang, & Koen, 2010) is associated with the SEQL (Flowe, 2011; Flowe & Ebbesen, 2007; Gronlund, 2004), while the familiarity process (which reflects a more global measure of memory strength or stimulus recency - Yonelinas et al., 2010) is associated with the SIML (Flowe, 2011; Flowe & Ebbesen, 2007; Gronlund, 2004). Moreover, and consistent with the relative judgment decision strategy (Lindsay & Wells, 1985), the presence of the stimuli at the same time provides a contextual basis for the application of familiarity (Meissner et al., 2005).

Thus, the recollection process used in the SEQL (first phase), which promotes a lower number of false positives (Flowe & Ebbesen, 2007; Gronlund, 2005), was not entirely replaced by a familiarity process in the SIML (second phase), but rather mixed. In the second phase of procedure B, there seems to have been a maintenance of the recollection process, but the increased sense of familiarity with the presence of the target has made the participants more successful, achieving a greater number of **Hits**. Importantly, Mickes, Wais, & Wixted, (2009) showed that, unlike we are led to believe, the recollection is not a categorical process (occurs or does not occur). Instead, it is continuous, as is the familiarity (which comes in levels), making the dual process theory compatible with the signal detection theory.

Gaze Measures

Regarding the eye tracking measurements, one of the most salient results was that participants who viewed the SEQL-StopR had a higher dwell time on the faces of all the suspects (criminal and distractors) than those who performed the SIML (comparison of the first phase in Procedures A and B). This result replicates the effect found by Flowe (2011) and seems to suggest that the SEQL and its absolute judgment encourages a more detailed visual exploration than the SIML with its relative judgment. Fascinating was the fact that, when comparing the two lineups of procedure B, the SIML of the second phase did not obtain statistically significant differences in dwell time on the faces for the categories '**Misses**' (despite the existence of a trend) and '**False Alarms** when the **target was absent**'. In fact, in these response categories, no changes in behavioural parameters (**Hits** and **Errors**) were observed from the first to the second phase. These results support some of the explanations above, since the dwell time is not different between SEQL-StopR and SIML (B – After SEQL-StopR), which seems to indicate the conservation of the absolute judgment in the SIML, being verified a more detailed exploration. One more clue to this retention of judgment is that the '**Misses**' response pattern (more dwell time in the distractors than in the target) matches in the SEQL-StopR and SIML, being different for any of the other SIML (A and C) whose pattern obtained for the **Misses** is "greater dwell time in the target than in the distractor". This effect of the absolute judgment for the SIML becomes more evident and more interesting when comparing the **Misses** in the SIML procedures A, B and C. The SIML performed after the SEQL-StopR generates significantly longer dwell time on the faces of the suspects than the SIML procedure A. Interestingly, although the SIML of procedure C (after SEQL-Pass) provides less dwell time on the faces than the SIML B, the differences were only marginally significant, showing a dimensionality (continuum) of the influence of the absolute judgment on the SIML. In line with this, the greatest influence of the absolute judgment is recorded in the SIML after the SEQL-StopR. When we compare the 3 SIML (A, B and C), regarding the category of **false positive** responses when the **target is absent**, whose behaviourally results don't change from the first to the second phase of procedure 2, the findings support the hypothesis of the influence of the absolute judgment in the SIML after the SEQL-StopR. That is, the dwell time on the faces of the suspects in the SIML after SEQL-StopR was significantly higher than the dwell time on the faces of the suspects in the SEQL of procedure A. Again, despite the SEQL of procedure C (after SEQL-Pass) originating less dwell time on the face, the differences were not statistically significant, showing the dimensionality of the influence of the absolute judgment. Additionally, although no significant differences were recorded, the response pattern of group B participants in the **Hits** category in the SIML after SEQL-StopR (slightly more dwell time on the target face than on the mean of the distractors) is inverse to the pattern that they had in the same category in the SEQL-StopR (approximately one second more in the mean of the distractors than on the target face). This result seems to show that the relative judgment and the familiarity assessment of the SIML allowed the participants to look more towards the target, increasing the number of **Hits**.

Still, regarding dwell time, it was expected that the participants who viewed the SEQL had more dwell time on the external characteristics of the face than the participants who looked at the SIML. However, this only happened for the distractors in the '**False Alarms** in **Target Absent**', there is no more differences between the SEQL and the SIML (A and B). This result was in contradiction to that postulated by Flowe (2011).

However, these data may be justified by the fact that the results obtained by Flowe (2011) contemplated the use of SEQL without the stopping rule, where people visualized all the suspects. Indeed, the superiority of the internal face features, mainly of the eyes and mouth in both lineups types (SEQ and SIM) is also not new, since the literature in face perception shows that either in learning or test phases, analyses of eye movements have discovered that fixations are directed toward the internal regions of the face (Henderson, Williams, & Falk, 2005; Luria & Strauss, 1978), especially the eyes and nose (Janik, Wellens, Goldberg, & Dell'Osso, 1978; Stacey, Walker, & Underwood, 2005). Also, it is interesting to note that when comparing the SIML A and B, we found that in the **Misses** category, the participants who visualized the SIML after the SEQL-StopR had more dwell time in the external characteristics of the distractor face than the participants who performed the SIML A. For the category '**False Alarms** in **target absent**' the pattern of results was the same, however, in this case, the difference constituted only a statistical trend. This result, which shows differences in the facial processing of distractors in **Misses** and False Positives in Target Absent between SIML A and B, supports the hypothesis of the existence of an absolute judgment mixed with a relative judgment, caused by a carry-over effect. One of the interesting results of the present study was the fact that dwell time in the target in the SEQL-StopR was a predictor of success. In fact, for the SEQL, the dwell time in the target was significantly greater when the participants identified him than when they did not identify anyone (**Misses**), or when they identified an innocent (**False Alarms**). This result can be explained by the fact that, in the '**Hits**', the greater memory trace (Steblay et al., 2001), and a feeling of recollection (Flowe & Ebbesen, 2007; Gronlund, 2004) promote more dwell time in the target's face.

Regarding the visits/fixations, it was expected for the number of visits to be greater in the SIML that occurred after the SEQL-StopR. This result was verified, showing once again that the absolute strategy of the first phase makes the visual exploration more exhaustive in the second phase. However, despite the difference in absolute values, the difference was not statistically significant. This scarce outcome can be explained by the fact that the relative judgment is closely related to the comparisons between the suspects – thus leading to less fixation in one particular stimulus (Mansour & Flowe, 2010).

Also, about the **number of visits**, it was expected from the literature that in the target absent condition, the average number of visits to the unidentified distractors be greater when a correct identification is performed than when an innocent (false alarm) is identified. This result was replicated by our study in the three SIML, giving some reliability to this indicator found by Flowe and Cottrell (2011).

Another interesting finding was that in the classic SIML (procedure A) when participants hit the target (**HITS**), they gave significantly more visits to the target than to the average of the distractors. In addition, this result has not been reached by any other category of response, for the average number of visits to the distractors when the participants hit the target are significantly higher than when they make a false positive. This result can be explained by the type of judgment used in the SIML, i.e., when a participant is facing a set of faces that appear simultaneously and selects a face, the relative judgment causes it to enter a deletion process, comparing the other faces with the selected one (Dunning & Stern, 1994; target→distractor 1→target→distractor 2→target→... until decision). Thus, the selected face ends the lineup with a higher number of fixations/visits than the average of the distractors. Interestingly, and crossing this result with the dwell time in the category of **Hits** for the SIML of procedure A, significantly longer dwell times were recorded in the eyes of the target than in the average of the eyes of the distractors. According to some literature, looking at the internal characteristics of the face, particularly the eyes, suggests that a familiarity-based process is in place (Olivares & Iglesias, 2008). This is a positive thing when we look at the eyes of a target and we judge it to be familiar, but it is harmful when we look into the eyes of an innocent and judge him as familiar. In fact, this situation occurred in our study, since for the SIML (A), when the target was absent, the dwell time in the eyes of the identified was significantly higher than in the eyes of the unidentified distractor. This realization of familiarity might have caused the participant to make a false alarm.

In summary, the present study has provided some empirical evidence based on behavioural and eye tracking results, that a procedure consisting of a SEQL-StopR in the first phase and an SIML in the second, takes advantage of the strengths of the two procedures. Thereby generating less **False Positives** in the SEQL, and being more successful in identifying the criminal in the SIML, without originating less False Positives. We also replicated some data already advanced by other investigators as to whether absolute judgment prompts more detailed visual exploration (Flowe, 2011), or if the average number of visits to unidentified distractors is greater when a correct identification is performed than when an innocent is identified (Flowe & Cottrell, 2011). The present study reaches similar conclusions as for the previous eye tracking studies (Flowe, 2011; Flowe & Cottrell, 2011; Malpass, Tredoux, & McQuiston-Surrett, 2009; Mansour & Flowe, 2010) regarding the use of the variables extracted from eye tracking as potential Hit markers. However, these variables still seem very fragile in predicting accuracy and will have to be the target of further studies. Probably, variables such as pupil dilation, closely related to attentional mechanisms (Kang, Huffer, & Wheatley, 2014) will have more promising results.

Although promising, the results related to the fusion of the lineups have to be the target of replication. Some limitations that may weaken our results should be taken into account in subsequent studies. One of the problems of our study was that we used the mean of distractors to facilitate analysis. This procedure may conceal potential effects such as a particular distractor has more dwell time than the target, but the average of the distractors is significantly lower.

Nevertheless, in our opinion, the results of the present study are a valuable increment in the literature of the eyewitness testimony and show a potential paradigm shift in the long-living discussion of which is the best lineup method.

CHAPTER IV

CENTRAL AND PERIPHERAL PATTERNS IN EYEWITNESS TESTIMONY

INTRODUCTION

The errors in eyewitness testimony have been a concern for over 100 years (e.g., Müstenberg, 1908). Indeed, there is a large body of findings that shows the fallibility of this evidence (e.g., Buckhout et al., 1974; Busey & Loftus, 2007; Clifford & Hollin, 1981), especially after the 1970s (Wells, 2014), when psychologists presented a scientific demonstration of these errors and reported them to the judicial system (Wells & Olson, 2003). This fact became more relevant after the advent of DNA evidence (Lacy & Stark, 2013), with exoneration studies gaining prominence in the eyewitness literature. The most well-known exoneration project is the Innocence Project, which celebrates 25 years this year. It indicates that, in 70% of the exonerations proved by the analysis of DNA, the conviction was made through an identification error of an eyewitness testimony (Innocence Project, 2017). Gross and Shaffer (2012), who led a joint project of Michigan's and Northwestern's law schools, analyzed, in detail, 873 cases in the National Registry of Exonerations and determined that 76% of these cases (667) be determined by eyewitnesses misidentifications. Smith and Cutler (2013) analyzed 1198 cases of wrongful convictions and found that in about 50% of the cases mistaken identifications were involved. This type of identification errors, called false positives, is the most problematic as it implicates an innocent person being sent to prison (Malpass, 2006).

These eyewitness errors have been diminished by the use of lineups (procedures where suspects are placed close to people who look like them), but the number of witness errors remains very large (Wells, 2014). The causes of errors in eyewitness testimony have been the subject of extensive research, and if factors such as system variables [e.g., lineup administrator (Greathouse & Kovera, 2009; Harris & Rosenthal, 1985), lineup size (Levi, 2007; Wells, Small, & Penrod, 1998), instructions (Malpass & Devine, 1981; Wise, Cushman, & Safer, 2012)] can be controlled by imposing evidence-based methodologies, factors relating to eyewitness characteristics [e.g., age (Pozzulo, Dempsey, Crescini, & Lemieux, 2009; Wise, Dauphinais, & Safer, 2007), emotional status (Christianson, 1992; Forgas, Laham, & Vargas, 2005), cognitive disabilities (Manzanero, Contreras, Recio, Alemany, & Martorell, 2012; Ternes & Yuille, 2008)] or to the characteristics of the event [delay (Kassin, Tubb, Hosch, & Memon, 2001; Penrod, Loftus, & Winkler, 1982), stress (Brigham, 1991; Deffenbacher, Bornstein, Penrod, & McGorty, 2004), time of exposure (Meissner, Sporer, & Schooler, 2007)] are very hard to control (Narby, Cutler, & Penrod, 1996). In his paper in Psychological Science, Wells (2014), one of the foremost researchers in the field of eyewitness testimony, referred to themes such as the probative value of the eyewitness, and stated in his concluding remarks that

"…a wholly different approach probably would be developed, perhaps one involving eye movements, pupil dilation, event-related potential patterns, response latencies, implicit memory tests, and other potential indicia of recognition. Bringing psychological science to bear on the serious problem of eyewitness identification ought to mean much more than manipulating whether photos are shown as groups versus one at a time. The next generation of eyewitness researchers should throw out the traditional" (p. 15).

These new types of measures, in particular, those looking for psychophysiological markers of recognition in eyewitness paradigms, have seldom been studied. There are a few studies that focus on eye movement patterns, examining decision-making processes in lineup using eyetracking (e.g., Flowe, 2011; Flowe & Cottrell, 2010; Mansour, Lindsay, Brewer, & Munhall, 2009). Nonetheless, there is little scientific evidence, and all of these researchers are cautious in generalizing the results. Regarding the event-related potentials (ERPs), as far as we are aware, only one study investigated the impact of identification accuracy on the brain's electrical response. In this study, developed by Lefebvre et al. (2007), the P300 remained a reliable predictor of correct identifications. The theoretical assumption behind this finding is generalized from the lie-detection literature, which shows that the P300 (positive potential appearing 300 ms after the stimulus onset) is elicited with greater amplitude when we visualize a relevant stimulus (usually taken from the crime scene). This differential elicitation only occurs for individuals with knowledge of the stimuli from the crime environment (Allen & Iacono, 1997; Farwell & Donchin, 1991; Rosenfeld, 2002). Adding to this explanation, the result obtained by Lefevre and colleagues (2007) can be framed within the theory of emotional processing (Lifshitz, 1966), which advocates that high arousal images (negative and positive affect) elicit a greater amplitude of the P300 than neutral images that are low arousal (e.g., Olofssen & Polisch, 2007, Deplaque et al, 2005). In fact, the critical item, being emotional/expected to induce high arousal, does elicit a larger amplitude of the P300 component (see Righi, 2012).

Although there are no more studies that directly relate ERPs or neuronal patterns with eyewitness testimony accuracy, there is a great body of literature that shows the psychophysiological distinction between familiar (perpetrator) and unfamiliar (distractor) faces, or the specific psychophysiological response to unpleasantness/arousal/negative affect (as the perpetrator) vs. neutral faces (as distractors) (see Werner, Kuhnel, & Markowitsch, 2013).

Literature has shown the sensitivity of the P100 (positive peak around 100 ms after stimulus onset) to face processing (e.g., Herrmann et al., 2005, Cunningham et al., 2012). A study investigating the effect of emotional facial expressions in recognition memory showed greater amplitude of the P100 for neutral faces that were presented in encoding with a fearful expression, compared with neutral faces learned/previously presented with a neutral expression and with new faces (Righi et al., 2012). The orbitofrontal region was pointed out as the producer of this difference. The authors postulate that this result is due to an association between the P100 and attentional capture by fear or threat stimuli (see Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007). Indeed, emotionally activating stimuli elicit a greater amplitude of the P100 than neutral stimuli (e.g., Carritié et al., 2004; Hot et al., 2006). Interestingly, Liu, Harris, and Kanwisher (2002), identified a face-selective response at approximately 100 msec that is correlated with correct face recognition. In addition to the P100, a visual sensory component of face processing (Bruce & Young, 1986), also the N170 (Peak with negative amplitude around 170 ms after stimulus onset), which is a structural processing index of the face (Bruce & Young, 1986), has been shown to be sensitive to emotional facial expressions (e.g., Caharel, DArripe, 2006). However, while some studies have shown a smaller (e.g., Blau, 2007) or greater (e.g., Righi, 2012) amplitude of the N170 for emotionally activating faces, conversely, other studies, show that the N170 is not affected by emotional aspects such as facial expressions (see Rossion and Jaques, 2008). Remarkably, some studies have shown the sensitivity of the N170 to familiarity (Caharel et al., 2002; Barret et al., 1988), yet other studies show that the N170 does not distinguish familiarity in faces (Jemel et al., 2003; Bentin et al., 1996). Herrmann et al. (2005), revealed that the N170 appears to be elicited in the inferior occipital cortex, in the fusiform gyrus, replicating the results found by Shibata et al. (2002). In this study, Herrmann et al. also found a hyperactivation in the medial prefrontal cortex and anterior cingulate cortex.

Regarding hemodynamic measures and referring to the familiar/unfamiliar face processing paradigm, Gobbini and Haxby (2006) found higher activation in the precuneus while watching familiar faces. Observing new faces led to higher responses in the fusiform gyrus and the amygdala. Von Der Heide et al. (2013) found higher left-lateralized anterior temporal lobe activations for familiar faces and a right anterior temporal lobe activation for unfamiliar faces. Concerning the emotion processing paradigm, faces with emotional cues led to a hyperactivation of the amygdala when compared to neutral faces (Adams et al., 2003, Adolphs, 2008). An amygdala–hippocampal cluster activation was found in the perception of fearful faces (Phillips et al., 2004. A study conducted by Reinders et al., (2006) established that the latency of the BOLD (blood-oxygen-level dependent) activity in the amygdala-hippocampal cluster (bilaterally) was found to be more than 500ms earlier for fearful faces than for neutral faces.

Regarding peripheral measures, as far as we know, there are no studies that relate these physiological responses to eyewitness testimony. However, we know that emotion and stimuli significance (positive or negative affect) play an important role in cardiac response (e.g., Bradley, 2009; Lang et al., 1997) and in facial muscle activity (e.g., Dimberg, 1982; Zhag et al., 2010). Regarding heart rate, Campbell et al. (1997) postulate a deceleration in heart rate when the organism is confronted with threatening cues. In this sense, the literature has shown that cardiac deceleration is significantly higher for unpleasant images compared to neutral or even pleasant ones (e.g., Bradley and Lang, 2000; Bradley, 2009). Heart Rate acceleration occurred when people are facing happy faces and deceleration when they are facing angry faces (e.g., Johnsen, Thayer, & Hugdahl, 1995; Palomba, Angrilli Mini, 1997).

Regarding electromyography, Schwartz and his colleagues found that, whereas unpleasant imagery elicited greater activity over the corrugator supercilii, pleasant imagery elicited higher activity over the zygomaticus major (Brown & Schwartz, 1980; Schwartz, Fair, Salt, Mandel, & Klerman, 1976). The activity over the Corrugator Supercilii is inversely related to the valence of a subjective experience (Larsen, Norris and Cacciopo, 2003; Kawamoto, Nittono, Ura, 2013), being associated with negative affect (Cacioppo, Martzke, Petty, & Tassinary, 1988; Dimberg, 1982) and representing a motivation to withdraw (see Davidson, 1990) from a stimulus (Allen, Harmon-Jones, & Cavender, 2001). In turn, zygomatic activity is directly associated with positive valence (Ribeiro et al., 2007) and related to positive affect (Cacioppo et al., 1988), which is associated with a motivation to approach (Davidson, 1990).

This study aimed to explore the psychophysiological markers of the different performances in eyewitness testimony. Simultaneously, and according to the reviewed literature, it was expected that the visualization of the author of a crime would elicit a greater amplitude of the P300 and the P100, greater activation of the corrugator supercilii and a deceleration in heart rate (due to the unpleasant nature of the author of a crime). We also aimed to clarify the role of the N170 in eyewitness recognition, as well as explore the zygomaticus major response in an eye-witness paradigm.

METHOD

Participants

Fourty undergraduate right-handed students (21 women), aged 18 to 26 years (*M* = 21.3; *SD* = 2.83), were recruited from the University of Aveiro.

Of these 40, only 29 participants obtained 20 valid epochs for each experimental condition, after signal pre-processing and epoching (see the following sections). Therefore, our final sample consisted of 29 participants (14 women), with a mean age of 21.9 years (*SD* = 2.41; range 18-26).

Stimuli and Task

Videos and Photos

The eight previously recorded theft videos (as described in Chapter II) were used in this experiment. In all these videos the face of the culprit was presented in frontal view during 4 of the 20 seconds of the full-length of the video. During the remaining time, the culprit was visible, but not in frontal view. All face images used in the subsequent lineups were grayscale, 10 cm height x 6 cm (On Screen Display) width, emotionally neutral and were rated as similar to the culprit of the video and with average attractiveness and distinctiveness (for additional details on the selection of the faces for the lineups, please see Chapter II).

Lineup procedures

The lineup used in this experiment was the SEQL without stopping rule, consisting of 10 cycles of six faces each, where five distractors and the culprit seen in the video (it was always a target present condition), appeared sequentially, in random order.

Stimuli presentation and the recording of response (key presses) were controlled by E-Prime software (Psychology Software Tools, Pittsburgh, PA).

Procedure

Experimental task

The task was organised in eight experimental blocks. Each experimental block (see Figure 37) consisted of a scene of a theft at an ATM. After watching the video, participants waited 120 seconds and typed the code provided by the researcher on the computer (interference task). Finally, eight cycles of the corresponding sequential lineup with five distracting faces and a target were presented, and participants had to decide for each face whether the individual they were watching was the one whom they had seen in the video committing the crime. Their answers should be given by pressing the key “S” for “yes” or the key "N" for “no”.

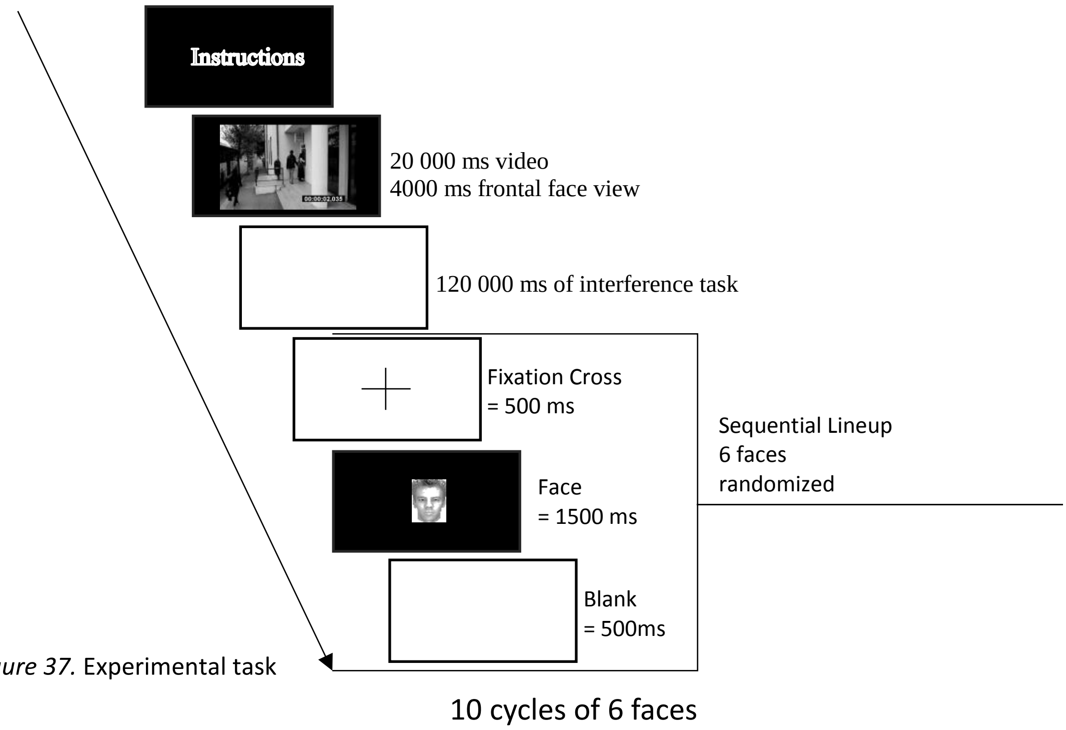


Figure 37. Experimental task

The written Consent form was signed before and after the participant performed the experimental task. This study was conducted in accordance with the Declaration of Helsinki.

Psychophysiological recordings and signal pre-processing

Central Measures

EEG activity was recorded from 32 electrodes mounted on a Waveguard Cap according to the 10–20 system and some intermediate positions. Recordings were made with a linked mastoid physical reference. The EEG was amplified with ANT® amplifiers with a gain of 30,000, recorded continuously at a sampling rate of 2048 Hz with ANT Eeprobe software and the impedance of all electrodes was maintained below 5k. The electrooculogram (EOG) was recorded from vertical EOG electrodes placed above and below the right eye and from Horizontal EOG electrodes placed in lateral to each eye.

Each continuous EEG underwent a sequence of preprocessing steps: i) ocular artifacts were rejected with the eye-movement correction algorithm used in the EEprobe (ANT, The Netherlands), ii) band-pass filtering at 0.1–30 Hz was used, iii) amplitude band-pass – artefact removal [-150 – 150μV], and iv) resulting data were re-referenced to the average of the left and right mastoid electrodes.

Peripheral Measures

Facial electromyography (EMG) and electrocardiography (ECG) signals were relayed through a shielded cable to Biopac amplifiers (Biopac Systems, Inc., Santa Barbara, CA), where signals were amplified with a gain of 5,000 and digitized at 2000 Hz.

The ECG signal was recorded using the D2 of Einthoven (positive electrode - left leg, negative electrode – right wrist and earth electrode - right leg). The EMG signals from the Corrugator Supercilii and the Zygomaticus major were recorded using two pairs (one pair for each muscle) of shielded electrodes, placed above the left eye and left cheek, respectively, according to the recommendations of Fridlund and Cacioppo (1986).

The raw EMG signals were transformed into integrated EMG, considering the area under the curve of the rectified EMG signal/the mathematical integral of the absolute value of the raw EMG signal.

Design and Analysis

The experimental design was a 2 x 2 within subjects. The first variable is the type of suspect (target or distractor) and the second variable is the accuracy of the identification (correct or incorrect). The intersection of these two variables results in 4 within-subjects experimental conditions, which theoretically match the Signal Detection Theory: i) Target Correct (Hit), ii) Target Incorrect (Miss), iii) Distractor Correct (Correct Rejection) and iv) Distractor Incorrect (False Alarm).

ERP Analysis

EEG signals were segmented into epochs of 900 ms, starting at 100 ms before stimulus onset and ending at 800 ms post-stimulus onset. Trials were baseline corrected from −100 ms to 0 ms and averaged by the experimental condition for each participant. Taking the literature into account, we analyzed the P100, N170 and P300 components and all time windows were defined in order to contain the respective peaks for each participant in all conditions. Specifically, the P100 was measured as the maximum peak positivity between 80 and 120 ms at occipital sites (O1, Oz, and O2). The N170 component was measured as the maximum peak negativity between 150 and 190 ms at occipitotemporal sites (P7 and P8). The P300 was measured in the temporal window between 300–600 ms and the mean amplitude in the 100 ms interval surrounding the peak latency was calculated for each participant and served as the dependent variable for P300 analyses. For this component, channels P3, PZ, P4, CP1, and CP2 were considered. All procedures were performed using the ASA lab software (ANT Software BV, Enschede, Netherlands).

Peripheral Analysis

Using the algorithm implemented in the Acknowledge software (Biopac), from the ECG graph (qrs graph), the heart rate was calculated in beats per minute. From the average heart rate obtained in the 1500 ms of after stimulus onset, we subtracted the average heart rate obtained in the 500 ms before the stimulus onset, just like other studies have done (Gamer et al., 2008). Following the same procedure, EMG reactivity was measured as the difference between activity during the 1500 ms post-stimulus onset and the 500 ms immediately prior to stimulus onset. Data averages per measure (ECG, EMGcorrutator, and EMGzygomaticus), experimental condition and participant were calculated using an algorithm purposely built in Matlab (Mathworks).

Source Analysis

Source analysis was performed with the software sLORETA (Standardized Low-Resolution Brain Electromagnetic Tomography (Pascual-Marqui, Esslen, Kochi, & Lehmann, 2002; Pascual-Marqui et al., 1999). The sLORETA algorithm is based on the statement that neighbouring voxels tend to activate synchronously. The sLORETA solution was computed using voxels that represent possible sources of the signal, which are restricted to the gray matter, based on the probabilistic brain tissue maps available from the Montreal Neurological Institute (MNI). Finally, the equally spaced grid points (5.00-mm grid spacing) and the recording array (32 electrodes) were superimposed on the Collins 27 MRI (Template T2) produced by the MNI. Log of F-Ratio comparisons was performed between the four experimental conditions on the periods that were significant in the ERP analysis. A single *t*-test was computed for averaged voxels in the time windows corresponding to the ERP component, based on 5000 SnPM randomizations with bulletproof, in order to correct critical thresholds and *p* values. The Talairach and MNI coordinates were used for labeling the corresponding brain areas. All analytical approaches were based on the software manual (download available in http://www.uzh.ch/keyinst/loreta).

Statistical Analysis

Data analyses were performed using SPSS and STATA. In what concerns the central measures, repeated measures 2 (Type: Target vs. Distractor) x 2 (Accuracy: Correct vs. Incorrect) x 3 (electrode: 2 or 3 or 5, depending on the component) ANOVAs were performed to verify the existence of effects of accuracy, stimulus type, electrode or interactions. Relatively to the peripheral analysis, repeated measures 2 (Type: Target vs. Distractor) x 2 (Accuracy: Correct vs. Incorrect) ANOVAs were performed to verify the existence of effects of accuracy, stimulus type or interactions. Paired samples *t*-tests were used to investigate the results from the ANOVAs in more detail. Due to the density of the results, the effect size was only calculated for the significant effects.

RESULTS

Behavioural Results

Accuracy analyses

Given the methodological restrictions inherent to ERP analysis in the present paradigm (i.e., elimination of cases with too many or too few errors, because they would not have enough epochs for analysis in each experimental condition), the error/hit ratios in the present study are artificially around 50/50 or 60/40. Therefore, descriptive and inferential analyses of accuracy would be meaningless and were not performed. It is important to note, however, that participants' performance that was analysed in the present work was not random. What happened was that some participants were able to identify correctly some of the targets (having close to 100% correct responses for that video/lineup) and were unable to recognise other targets (having close to 0% correct responses for that video/lineup). As data from all videos/lineups were analysed together, this gave an average performance around 50/60% correct, which was what was necessary to be able to have a minimum number of epochs to analyse in each experimental condition (as indicated in the Methods section).

Response times

Regarding response times, the average response times obtained in the four experimental conditions resulting from the crossing of the variable type (target vs. distractor) and accuracy (correct and incorrect response) are shown in Figure 38.

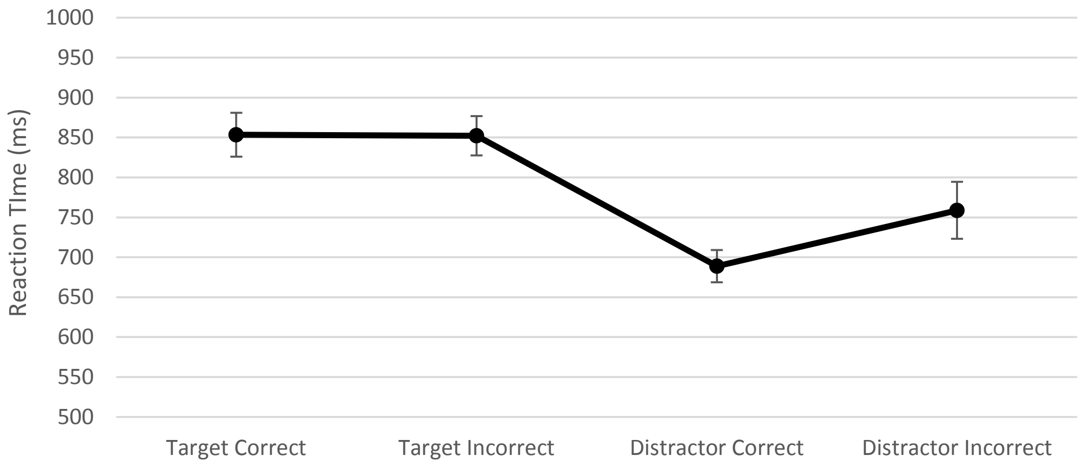


Figure 38. Average response times obtained in the four experimental conditions

By analyzing the absolute values (see Figure 38), we found greater response times for the targets than for the distractors and also greater response times for the incorrect answers than for the correct ones. Inferentially, the 2 within (type – target vs distractor) * 2 within (accuracy - correct vs incorrect) repeated measures ANOVA showed an effect of type, $F(1,28) = 31.991$, $p < .001$, $\eta_p^2 = .533$, and a statistical trend for accuracy, $F(1,28) = 3.557$, $p = .07$, $\eta_p^2 = .113$, but no interaction of type*accuracy, $F(1,28) = 2.786$, $p = .106$ was found.

Central Measures

Regarding the central measures, the results from ERP analyses are complemented with the analyses of the sLORETA for the corresponding time windows.

P100

Analyses of P100 focused on the locations where the P100 was maximal, namely channels O1, Oz and O2 (see Figure 39). Based on the averages of maximum amplitude in time window 80 – 120 msec, the 2 (type)* 2 (accuracy)* 3 (electrode) ANOVA showed an effect of accuracy $F(1,28) = 6.448$, $p < .01$, $\eta_p^2 = .187$, i.e., the amplitude of P100 was significantly higher in the incorrect responses than in the correct ones. Additionally, an effect of electrode $F(2,56) = 15.702$, $p < .001$, $\eta_p^2 = .359$ was observed, but no other main effect or interaction resulting from this ANOVA reached statistical significance or even a statistical trend.

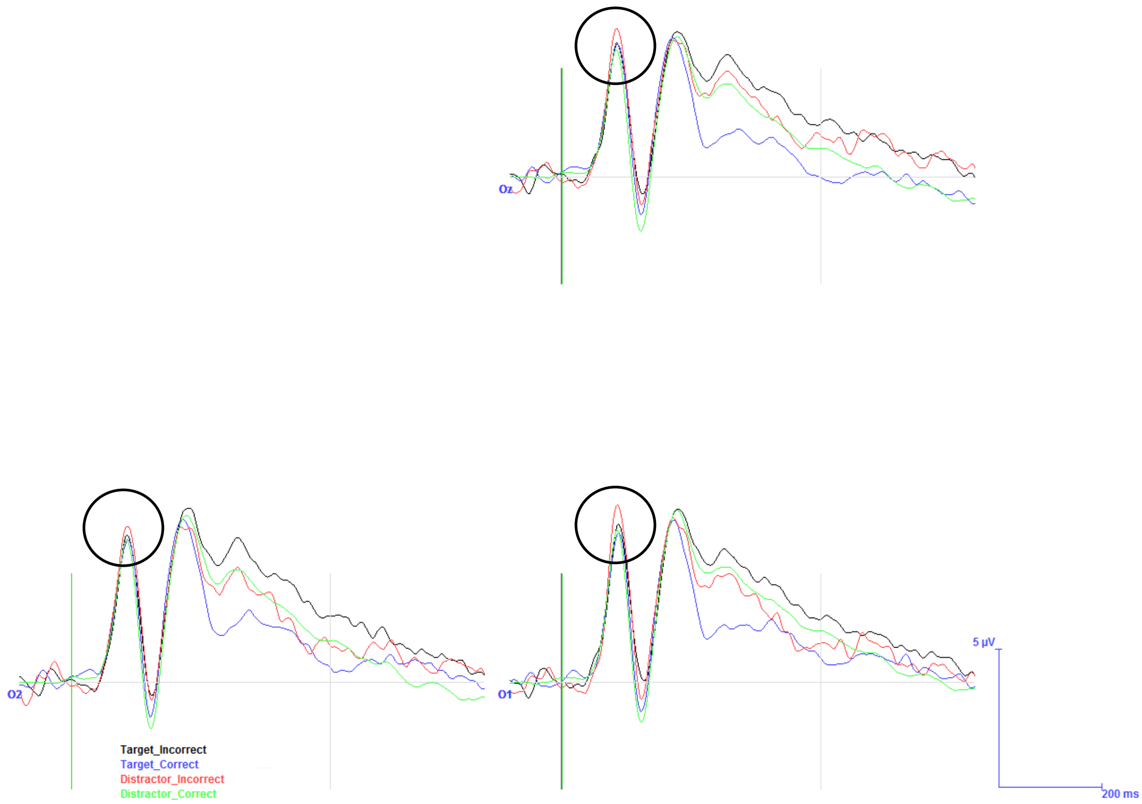
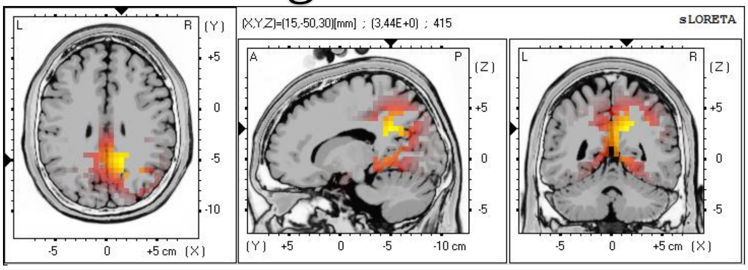


Figure 39. P100 in electrodes O1, OZ, and O2

Splitting the effect of type, two ANOVA's 2 (accuracy)*3 (electrode), one for each level (Target and Distractor) were performed. Relatively to Target, no effects of accuracy *F*(1,28) = .259, *p* = .615, or Interaction *F*(1,28) = .720, *p* = .481 were found. The main effect of electrode *F*(2,56) = 11.278, *p* < .001, η_p^2 = .238 was found for Targets, however, none of the three *t*-tests (one for each electrode) recorded significant results (all *ps* > .424). For the distractors effects of accuracy *F*(1,28) = 11.278, *p* < .001, η_p^2 = .287 and Electrode *F*(2,56) = 14.302, *p* < .001, η_p^2 = .338 were obtained. Due to the effect of the electrode, three *t*-tests - one for each electrode - were performed. Concerning the O1 channel, the amplitude of the P100 was significantly greater when the participants mistakenly identified a distractor (incorrect distractor) than when they correctly rejected the distractor (correct distractor), *t*(28) = 3.298, *p* < .01, 95% CI [0.37, 1.60], Cohens'*d* = .61, Cl effect Size = .73. For the OZ Channel, the amplitude of the P100 was again significantly greater when the participants mistakenly identified a distractor than when they correctly rejected the distractor *t*(28) = 2.898, *p* < .01, 95% CI [0.23, 1.35], Cohens'*d* = .547, Cl effect Size=.71. Regarding the O2, the results followed a similar pattern as in the O1 and OZ channels *t*(28) = 2.076, *p* = .047, 95% CI [0.006, 1.01], Cohens'*d* = .392, Cl effect Size = .65.

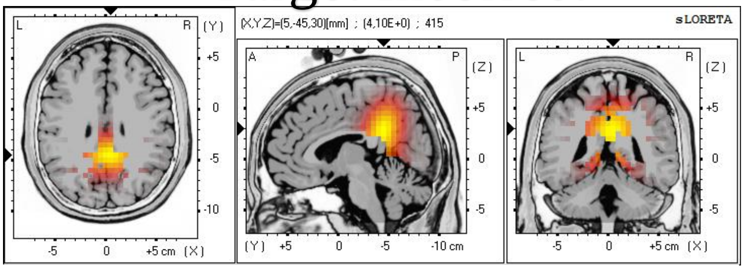
Due to its conceptual value, an additional comparison of the amplitude of the P100 component when an identification was made, either correctly or incorrectly, i.e., comparing the correct target (Hit) with the distractor incorrect (False Alarm), was performed. Indeed, for the electrodes O1 – *t*(28) = -4.103, *p* < .001, 95% CI [-1.49, -.500], Cohens'*d* = -.775, Cl effect Size = .219 and OZ – *t*(28) = -2.080, *p* = .047, 95% CI [-1.06, -.008], Cohens'*d* = -.393, Cl effect Size = .35, – the amplitude of the P100 was significantly greater when the participants mistakenly identified a distractor than when they identified the target correctly. Relatively to O2, although statistical significance was not achieved, the results showed a statistical trend $t(28) = -1.915$, $p = .067$, 95% CI [-1.14, -.038], Cohens'$d = -.393$, Cl effect Size = .34 pointing in the same direction. Considering the results obtained in the analysis of P100, temporal and spatial analyses were performed with sLORETA. The top 10 activation voxels (functional localization – MNI and Talairach coord.) in P100’s time window in each condition (target correct, target incorrect, distractor correct and distractor incorrect) are presented in Figure 40.

Target Correct



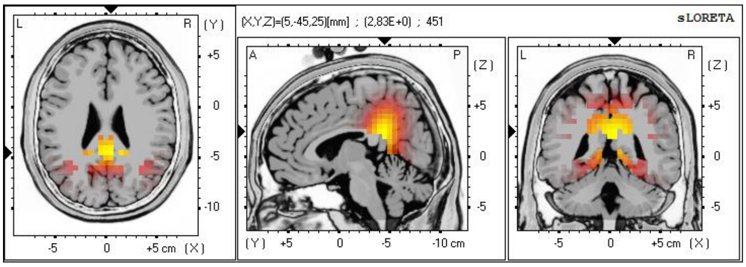
MNI			TALAIRACH						
X	Y	Z	X	Y	Z	Voxel Value	BA	Lobe	Structure
15	-50	30	15	-47	30	3.43660E+0000	31	Parietal Lobe	Precuneus
15	-55	30	15	-52	30	3.42518E+0000	31	Parietal Lobe	Precuneus
15	-45	25	15	-42	25	3.41655E+0000	31	Limbic Lobe	Cingulate G.
20	-50	35	20	-47	35	3.41381E+0000	31	Parietal Lobe	Precuneus
15	-50	35	15	-47	35	3.40683E+0000	31	Parietal Lobe	Precuneus
10	-50	30	10	-47	30	3.40592E+0000	31	Parietal Lobe	Precuneus
15	-45	30	15	-42	30	3.40443E+0000	31	Limbic Lobe	Cingulate G.
20	-45	25	20	-42	25	3.40327E+0000	31	Limbic Lobe	Cingulate G.
20	-45	30	20	-42	30	3.39713E+0000	31	Parietal Lobe	Precuneus
10	-55	30	10	-52	30	3.38548E+0000	31	Parietal Lobe	Precuneus

Target Incorrect



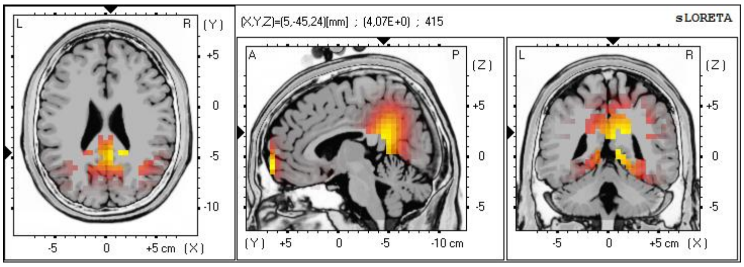
MNI			TALAIRACH						
X	Y	Z	X	Y	Z	Voxel Value	BA	Lobe	Structure
5	-45	30	5	-42	30	4.09814E+0000	31	Limbic Lobe	Cingulate Gyrus
5	-45	25	5	-42	25	4.09500E+0000	23	Limbic Lobe	Posterior Cing.
0	-45	25	0	-42	25	4.09079E+0000	31	Limbic Lobe	Cingulate Gyrus
0	-45	30	0	-42	30	4.09051E+0000	31	Limbic Lobe	Cingulate Gyrus
5	-45	35	5	-42	34	4.07751E+0000	31	Limbic Lobe	Cingulate Gyrus
10	-45	30	10	-42	30	4.07652E+0000	31	Limbic Lobe	Cingulate G.
5	-40	25	5	-38	25	4.07625E+0000	23	Limbic Lobe	Posterior Cing.
5	-45	20	5	-43	21	4.07280E+0000	30	Limbic Lobe	Posterior Cing.
0	-45	20	0	-43	21	4.07155E+0000	30	Limbic Lobe	Posterior Cing.
0	-40	25	0	-38	25	4.07056E+0000	31	Limbic Lobe	Cingulate Gyrus

Distractor Correct



MNI			TALAIRACH						
X	Y	Z	X	Y	Z	Voxel Value	BA	Lobe	Structure
5	-45	25	5	-42	25	2.83038E+0000	23	Limbic Lobe	Posterior Cing.
5	-45	20	5	-43	21	2.82864E+0000	30	Limbic Lobe	Posterior Cing.
0	-45	25	0	-42	25	2.82677E+0000	31	Limbic Lobe	Cingulate G.
0	-45	20	0	-43	21	2.82594E+0000	30	Limbic Lobe	Posterior Cing.
5	-40	25	5	-38	25	2.82300E+0000	23	Limbic Lobe	Posterior Cing.
0	-40	25	0	-38	25	2.82016E+0000	31	Limbic Lobe	Cingulate G.
5	-45	30	5	-42	30	2.81719E+0000	31	Limbic Lobe	Cingulate G.
0	-45	30	0	-42	30	2.81276E+0000	31	Limbic Lobe	Cingulate G.
-5	-45	25	-5	-42	25	2.81240E+0000	23	Limbic Lobe	Posterior Cing.
10	-45	30	10	-42	30	2.81057E+0000	31	Limbic Lobe	Cingulate G.

Distractor Incorrect



MNI			TALAIRACH						
X	Y	Z	X	Y	Z	Voxel Value	BA	Lobe	Structure
5	65	-10	5	63	-12	4.11406E+0000	11	Frontal Lobe	Superior Front. G.
15	-45	25	15	-42	25	4.10508E+0000	31	Limbic Lobe	Cingulate Gyrus
5	-45	20	5	-43	21	4.09522E+0000	30	Limbic Lobe	Posterior Cing.
15	-50	30	15	-47	30	4.09109E+0000	31	Parietal Lobe	Precuneus
5	65	-5	5	63	-7	4.08531E+0000	10	Frontal Lobe	Medial Front. G.
20	-45	25	20	-42	25	4.08498E+0000	31	Limbic Lobe	Cingulate G.
10	-50	30	10	-47	30	4.07934E+0000	31	Parietal Lobe	Precuneus
5	-50	15	5	-48	16	4.07902E+0000	30	Limbic Lobe	Posterior Cing.
10	-45	0	10	-44	2	4.07229E+0000	30	Limbic Lobe	Parahippocamp.G.
10	-40	0	10	-39	2	4.07127E+0000	30	Limbic Lobe	Parahippocamp.G.

Figure 40. Top 10 activation voxels (functional localization – MNI and Talairach coord.) in P100's time window in each condition (target correct, target incorrect, distractor correct and distractor incorrect).

Given the absolute differences verified in Figure 40 and taking into account the results obtained in the P100 ERP, with analyzes using ASALab, the following comparisons were made: Target Correct vs. Target Incorrect, Distractor Correct vs. Distractor Incorrect and Target Correct vs. Distractor Incorrect. Before estimating the functional location of the differences between these conditions, we checked in which latencies the *t* value reached the maximum value in the P100 time window. Much like in the analysis of ERP's in AsaLab only for Distractor Correct vs Distractor Incorrect comparison (Max t at 107 ms; $t > -4.948$, $p < .01$) and Target Correct vs Distractor Incorrect (Max t at 112 ms $t > 4{,}093$, $p < .01$), the maximum values of t were above the threshold to reach a significant difference. Using a procedure described by Pasqual-Marqui (2012), to evaluate the difference between Distractor Correct vs. Distractor Incorrect at latency 107ms, a log of the ratio of averages, SnPM randomizations, computed bulletproof, corrected critical thresholds and *P* values, and 5000 permutations were performed. Figure 41 contains a log of F-ratios for the one-tailed test: Distractor Correct < Distractor Incorrect at latency 107; and the difference of represented functional activation is significant at $p < .001$. The structures where significant differences were recorded (Voxel Value> -2.40), as well as their coordinates, are in Table 1.

Table 1. Brain structures where significant differences between Distractor Correct and Distractor Incorrect conditions were recorded – P100's time window

X(MNI)	Y(MNI)	Z(MNI)	X(TAL)	Y(TAL)	Z(TAL)	Voxel Value	BA	Lobe	Structure
5	65	-5	5	63	-7	-3.24939E+0000	10	Frontal Lobe	Medial Frontal Gyrus
5	65	-10	5	63	-12	-3.24286E+0000	11	Frontal Lobe	Superior Frontal Gyrus
10	65	-5	10	63	-7	-3.20780E+0000	10	Frontal Lobe	Superior Frontal Gyrus
10	65	-15	10	62	-16	-3.18421E+0000	11	Frontal Lobe	Medial Frontal Gyrus
35	60	-10	35	58	-11	-2.52454E+0000	10	Frontal Lobe	Middle Frontal Gyrus
-45	30	40	-45	31	35	-2.51966E+0000	9	Frontal Lobe	Medial Frontal Gyrus
-5	55	-25	-5	52	-24	-2.47973E+0000	11	Frontal Lobe	Rectal Gyrus
-35	40	40	-35	41	35	-2.40695E+0000	9	Frontal Lobe	Middle Frontal Gyrus

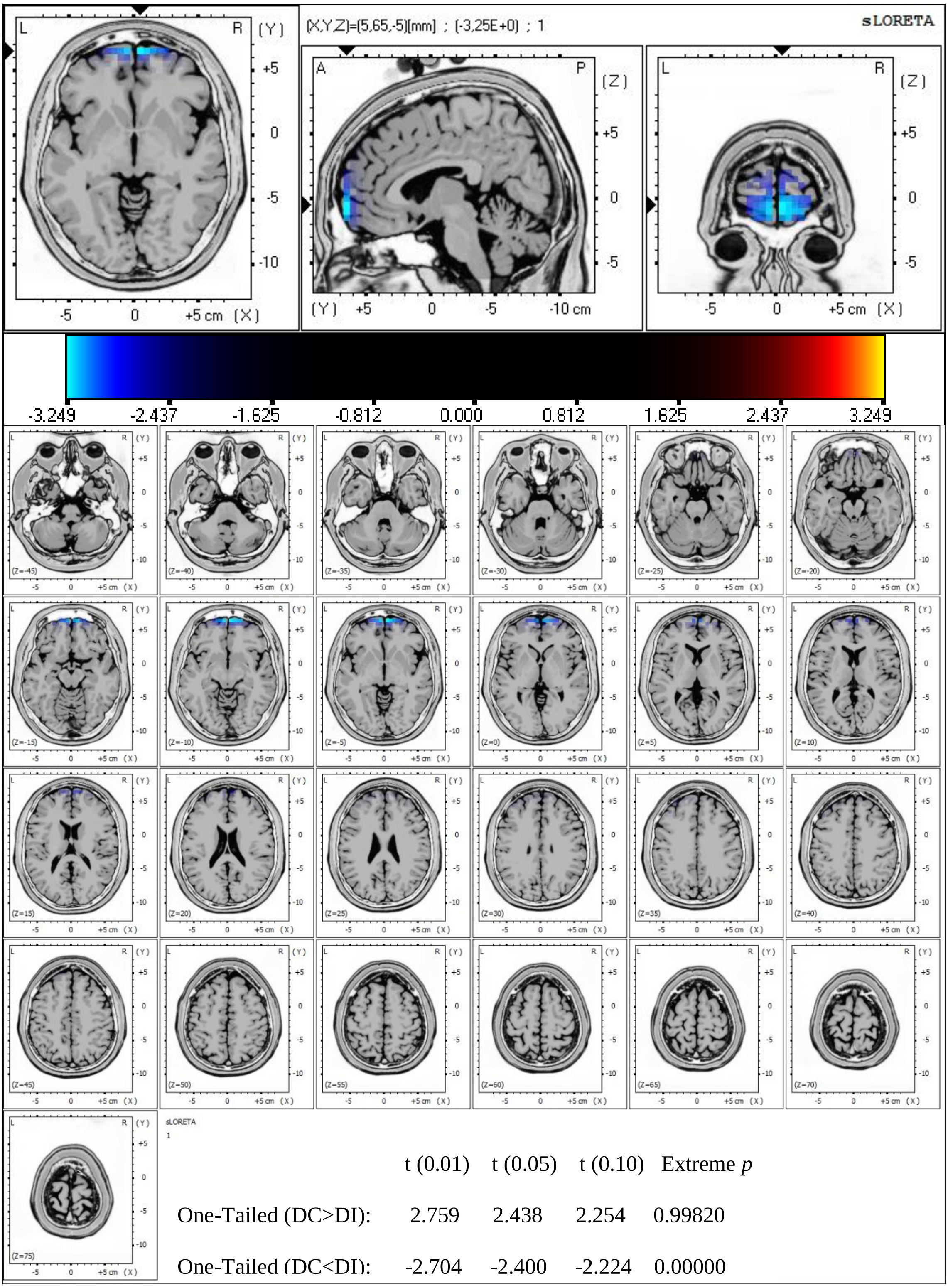
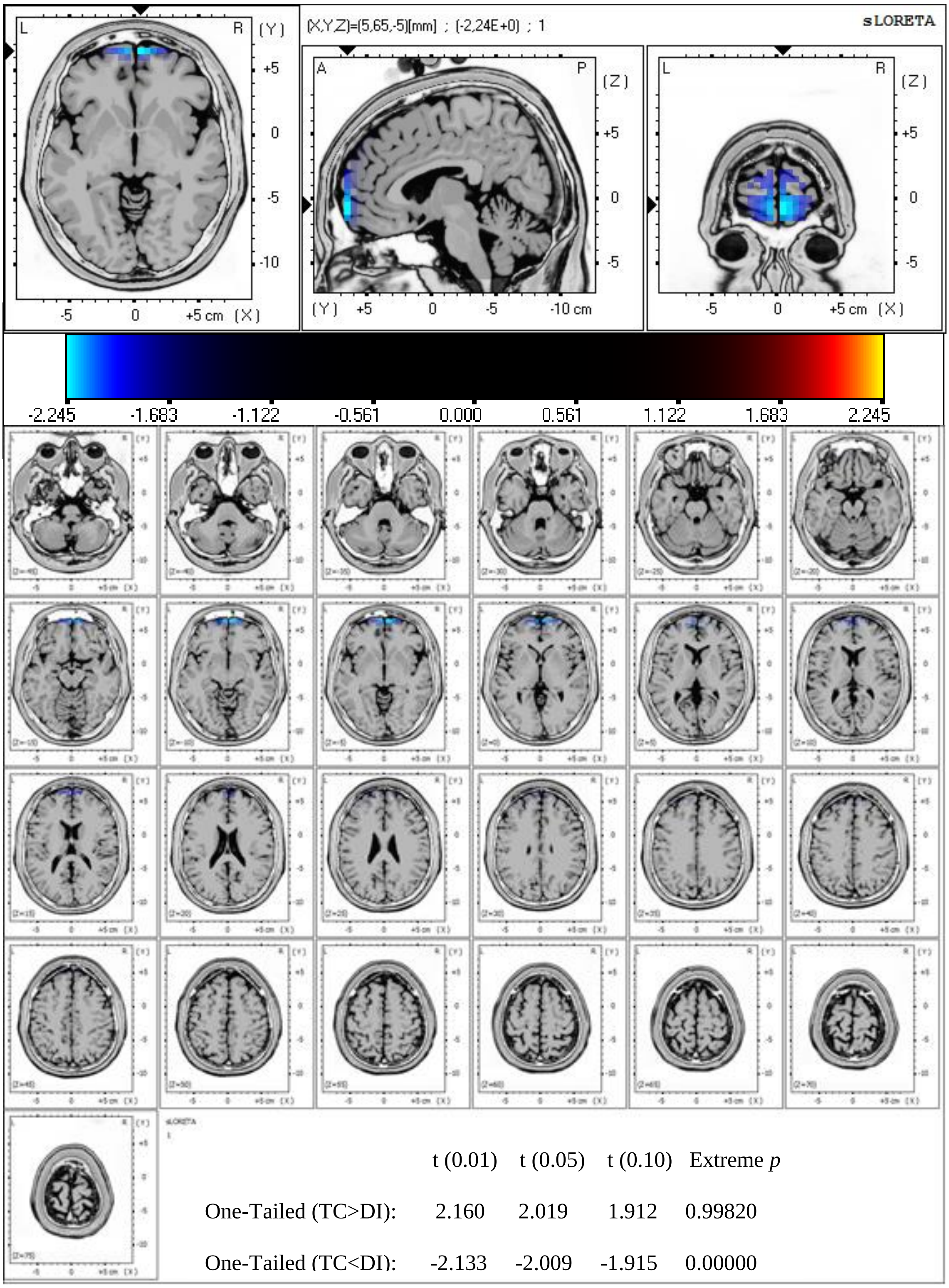


Figure 41. Log of F-ratios for the one-tailed test: Distractor Correct < Distractor Incorrect at 107 ms.

To explore the difference between Target Correct vs. Distractor Incorrect at latency 112 ms, the procedure recommended by Pasqual-Marqui (2012) was used once again. Figure 42 contains a log of F-ratios for the one-tailed test: Target Correct < Distractor Incorrect at latency 112; and the difference of represented functional activation is significant at p <.001. The structures where significant differences were recorded (Voxel Value> -2.009) as well as their coordinates are in Table 2.

Table 2. Brain structures where significant differences between Target Correct and Distractor Incorrect conditions were recorded – P100's time window

X(MNI)	Y(MNI)	Z(MNI)	X(TAL)	Y(TAL)	Z(TAL)	Voxel Value	BA	Lobe	Structure
5	65	-5	5	63	-7	-2.24462E+0000	10	Frontal Lobe	Medial Frontal Gyrus
5	65	-10	5	63	-12	-2.24449E+0000	11	Frontal Lobe	Superior Frontal Gyrus
10	65	-5	10	63	-7	-2.21472E+0000	10	Frontal Lobe	Superior Frontal Gyrus
5	65	0	5	63	-16	-2.18948E+0000	11	Frontal Lobe	Medial Frontal Gyrus



	t (0.01)	t (0.05)	t (0.10)	Extreme *p*
One-Tailed (TC>DI):	2.160	2.019	1.912	0.99820
One-Tailed (TC<DI):	-2.133	-2.009	-1.915	0.00000

Figure 42. Log of F-ratios for the one-tailed test: Target Correct <Distractor Incorrect at 112 ms.

N170

The ERP Analyses for this component were focused on channels P7 and P8 (see Figure 43), and all the minimum values obtained by the participants in the four conditions, were in the temporal window 150-190 ms. The 2 (type)* 2 (accuracy)* 2 (electrode) ANOVA showed an effect of Accuracy $F(1,28) = 8.848$, $p <.001$, $\eta_p^2 =.240$ and no other effect or interaction reached statistical significance. This means that the amplitude of the N170 is significantly lower for the correct responses than for the incorrect responses.

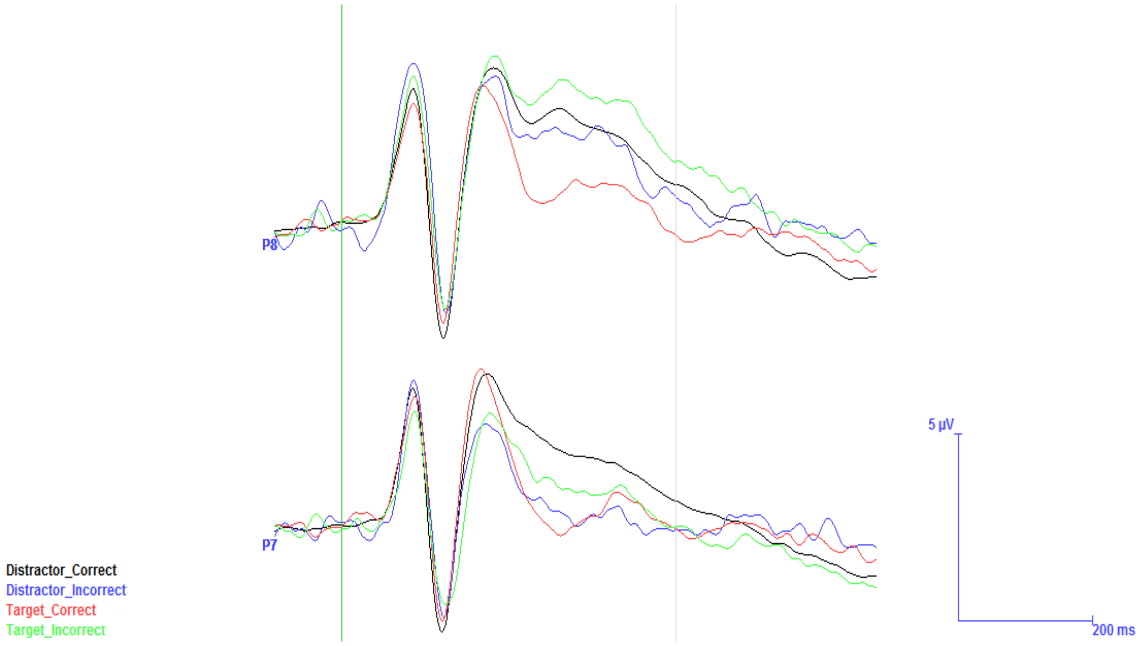


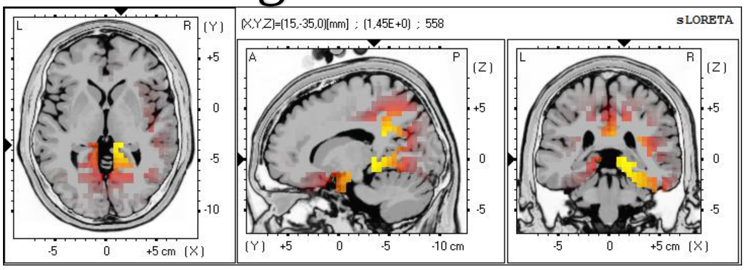
Figure 43. N170 in electrodes P7 and P8.

Two ANOVA's 2 (accuracy)*2 (electrode), one for Target and other for Distractor were performed. Relatively to Target, no effects of accuracy $F(1,28) = 2.722$, $p = .110$, electrode $F(1,28) = 0.175$, $p =.679$, or Interaction $F(1,28) = 0.002$, $p = .961$ were found. For the distractors, an effect of accuracy was obtained, $F(1,28) = 13.628$, $p <.001$, $\eta_p^2 = .327$, and no effect of electrode $F(1,28) = 0.007$, $p = .935$ or interaction $F(1,28) = 0.878$, $p=.357$ were obtained. The comparison "target correct vs distractor Incorrect", showed no significant differences (P7 and P8 *ps* > .374). Figure 44 shows the minimum average activations in each condition in the time window of the N170. The comparisons "Target Correct vs. Target Incorrect", "Distractor Correct vs. Distractor Incorrect" and "Target Correct vs. Distractor Incorrect" were performed. Before estimating the functional location of the differences between these conditions, we checked in which of the N170 time window latencies the *t*-test value reached the minimum value. Only for Distractor Correct vs. Distractor Incorrect comparison (Min t at 176 ms; $t > -3.061$, $p < .01$) the minimum values of t were above the threshold to reach a significant difference. Log of the ratio of averages, SnPM randomizations, computed bullet proof, corrected critical thresholds and *P* values, and 5000 permutations were performed. Figure 45 contains a log of F-ratios for the one-tailed test: Distractor Correct < Distractor Incorrect at latency 176; and the difference of represented functional activation is significant at $p < .001$. The structures where significant differences were recorded (Voxel Value > -1.574), as well as their coordinates, are in Table 3.

Table 3. Brain structures where significant differences between Distractor Correct and Distractor Incorrect conditions were recorded – N170's time window

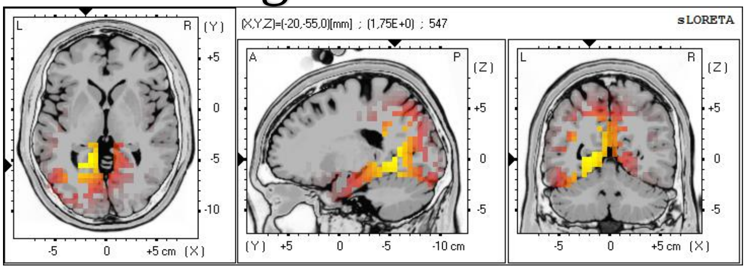
X(MNI)	Y(MNI)	Z(MNI)	X(TAL)	Y(TAL)	Z(TAL)	Voxel Value	BA	Lobe	Structure
-5	65	-10	-5	63	-12	-2.47206E+0000	11	Frontal Lobe	Superior Frontal Gyrus
10	65	-15	10	62	-16	-2.44243E+0000	11	Frontal Lobe	Medial Frontal Gyrus
-5	65	-5	-5	63	-7	-2.44165E+0000	10	Frontal Lobe	Superior Frontal Gyrus
5	65	-5	5	63	-7	-2.43800E+0000	10	Frontal Lobe	Medial Frontal Gyrus
-30	50	35	-30	50	30	-2.12281E+0000	9	Frontal Lobe	Superior Frontal Gyrus
-5	55	-25	-5	52	-24	-2.10220E+0000	11	Frontal Lobe	Rectal Gyrus
-30	60	5	-30	58	2	-2.03200E+0000	10	Frontal Lobe	Middle Frontal Gyrus
-10	55	-20	-10	52	-19	-2.01222E+0000	11	Frontal Lobe	Orbital Gyrus
5	55	0	5	53	-3	-1.98785E+0000	10	Limbic Lobe	Anterior Cingulate
-45	40	25	-45	40	21	-1.94886E+0000	46	Frontal Lobe	Middle Frontal Gyrus
-5	50	0	-5	48	-2	-1.83352E+0000	32	Limbic Lobe	Anterior Cingulate
-40	45	0	-40	44	-2	-1.68150E+0000	10	Frontal Lobe	Sub-Gyral
-45	45	0	-45	44	-2	-1.65900E+0000	10	Frontal Lobe	Inferior Frontal Gyrus
50	40	15	50	39	12	-1.63366E+0000	46	Frontal Lobe	Inferior Frontal Gyrus

Target Correct



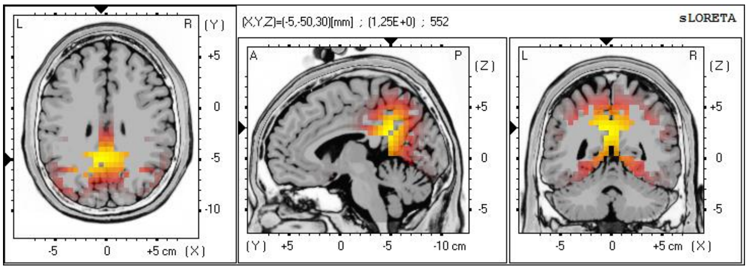
MNI X	MNI Y	MNI Z	TALAIRACH X	TALAIRACH Y	TALAIRACH Z	Voxel Value	BA	Lobe	Structure
15	-35	0	15	-34	2	1.44684E+0000	27	Limbic Lobe	Parahippoc. G.
20	-35	-5	20	-34	-2	1.44445E+0000	27	Limbic Lobe	Parahippoc. G.
15	-35	-5	15	-34	-2	1.44414E+0000	27	Limbic Lobe	Parahippoc. G.
20	-30	-5	20	-29	-3	1.44313E+0000	27	Limbic Lobe	Parahippoc. G.
20	-30	-10	20	-29	-7	1.44299E+0000	28	Limbic Lobe	Parahippoc. G.
20	-35	-10	20	-34	-7	1.44134E+0000	35	Limbic Lobe	Parahippoc. G.
20	-30	-15	20	-30	-11	1.44118E+0000	35	Limbic Lobe	Parahippoc. G.
15	-45	25	15	-42	25	1.44020E+0000	31	Limbic Lobe	Cingulate G.
15	-35	-10	15	-34	-7	1.44006E+0000	30	Limbic Lobe	Parahippoc. G.
20	-25	-15	20	-25	-11	1.43934E+0000	35	Limbic Lobe	Parahippoc. G.

Target Incorrect



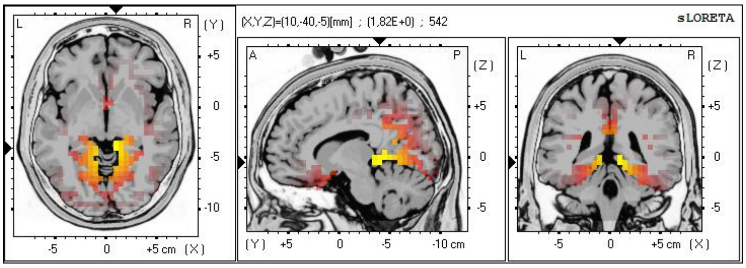
MNI X	MNI Y	MNI Z	TALAIRACH X	TALAIRACH Y	TALAIRACH Z	Voxel Value	BA	Lobe	Structure
-20	-55	0	-20	-53	3	1.74971E+0000	18	Occipital Lobe	Lingual Gyrus
-15	-50	0	-15	-48	2	1.74702E+0000	19	Occipital Lobe	Lingual Gyrus
-15	-55	5	-15	-53	7	1.74639E+0000	30	Limbic Lobe	Posterior Cing.
-25	-55	0	-25	-53	3	1.74483E+0000	30	Limbic Lobe	Parahippoc. G.
-20	-50	-5	-20	-49	-2	1.74369E+0000	19	Limbic Lobe	Parahippoc. G.
-15	-45	0	-15	-44	2	1.74340E+0000	30	Limbic Lobe	Parahippoc. G.
-20	-60	10	-20	-58	12	1.74131E+0000	30	Limbic Lobe	Posterior Cing.
-20	-55	-5	-20	-53	-2	1.74081E+0000	19	Limbic Lobe	Parahippoc. G.
-15	-55	0	-15	-53	3	1.74031E+0000	18	Occipital Lobe	Lingual Gyrus
-10	-50	5	-10	-48	7	1.73947E+0000	29	Limbic Lobe	Posterior Cing.

Distractor Correct



MNI X	MNI Y	MNI Z	TALAIRACH X	TALAIRACH Y	TALAIRACH Z	Voxel Value	BA	Lobe	Structure
-5	-50	30	-5	-47	30	1.25359E+0000	31	Parietal Lobe	Precuneus
-5	-45	25	-5	-42	25	1.25301E+0000	23	Limbic Lobe	Posterior Cing.
0	-45	25	0	-42	25	1.25279E+0000	31	Limbic Lobe	Cingulate G.
-10	-50	30	-10	-47	30	1.25268E+0000	31	Parietal Lobe	Precuneus
0	-45	20	0	-43	21	1.25110E+0000	30	Limbic Lobe	Posterior Cing.
0	-50	25	0	-47	25	1.25106E+0000	23	Limbic Lobe	Posterior Cing.
0	-50	30	0	-47	30	1.25098E+0000	31	Parietal Lobe	Precuneus
5	-45	25	5	-42	25	1.24873E+0000	23	Limbic Lobe	Posterior Cing.
-5	-45	30	-5	-42	30	1.24852E+0000	31	Limbic Lobe	Cingulate G.
0	-45	30	0	-42	30	1.24805E+0000	31	Limbic Lobe	Cingulate G.

Distractor Incorrect



MNI X	MNI Y	MNI Z	TALAIRACH X	TALAIRACH Y	TALAIRACH Z	Voxel Value	BA	Lobe	Structure
10	-40	-5	10	-39	-2	1.82465E+0000	30	Limbic Lobe	Parahippoc. G.
10	-40	0	10	-39	2	1.82151E+0000	30	Limbic Lobe	Parahippoc. G.
10	-45	0	10	-44	2	1.82150E+0000	30	Limbic Lobe	Parahippoc. G.
5	-50	5	5	-48	7	1.82126E+0000	29	Limbic Lobe	Posterior Cing.
-5	-50	5	-5	-48	7	1.81985E+0000	29	Limbic Lobe	Posterior Cing.
10	-35	-5	10	-34	-2	1.81983E+0000	27	Limbic Lobe	Parahippoc. G.
10	-45	5	10	-43	7	1.81510E+0000	29	Limbic Lobe	Posterior Cing.
10	-35	0	10	-34	2	1.81489E+0000	27	Limbic Lobe	Parahippoc. G.
10	-50	0	10	-48	2	1.81409E+0000	30	Limbic Lobe	Parahippoc. G.
10	-40	-5	10	-39	-2	1.82465E+0000	30	Limbic Lobe	Parahippoc. G.

Figure 44. Top 10 activation voxels (functional localization – MNI and Talairach coord.) in N170's time window in each condition (target correct, target incorrect, distractor correct and distractor incorrect).

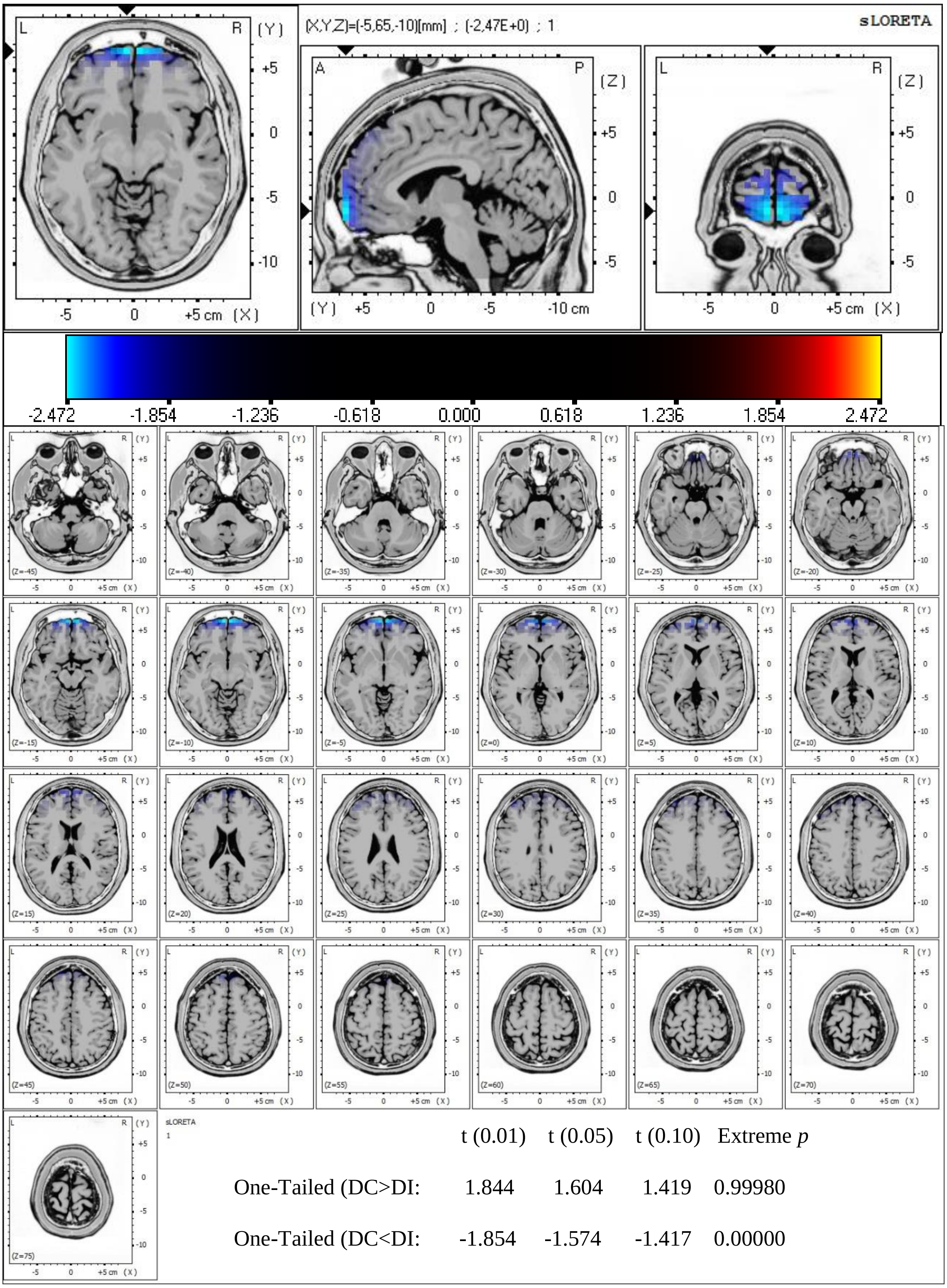


Figure 45. Log of F-ratios for the one-tailed test: Distractor Correct <Distractor Incorrect at 176 ms

P300

For P300 analyses, the channels P3, PZ, P4, CP1, and CP2 were taken into account (see Figure 46). The temporal window of 300–600ms was selected because it encapsulated the peak of the P300 for each of the participants in all conditions. The data showed higher amplitudes of the P300 for the target correct and for the distractor incorrect conditions, that is, whenever the participants made an identification. In fact, the 2 (type)* 2 (accuracy)* 5 (electrode) ANOVA showed an effect of Electrode $F(4,28)=9.695$, $p<.001$, $\eta_p^2 =.257$ and an interaction of type*accuracy $F(4,112) = 17.866$, $p < .001$, $\eta_p^2 =.390$. No other effects or interactions were found.

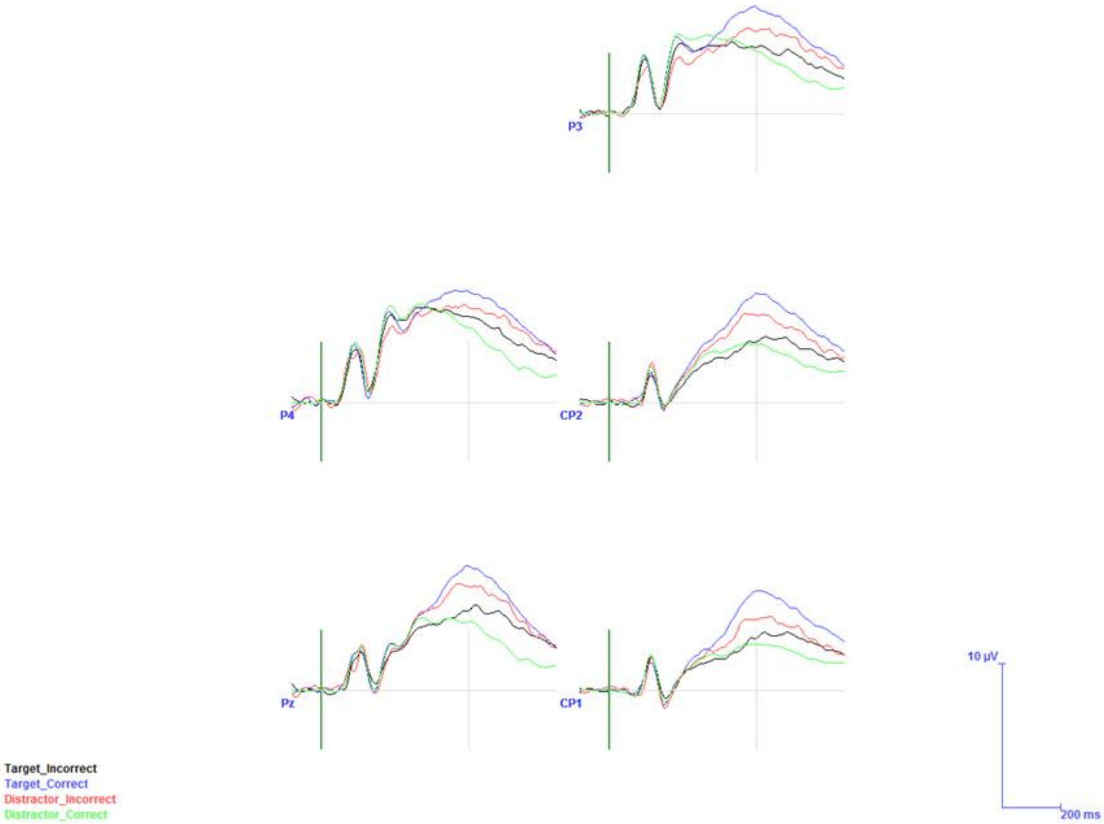


Figure 46. P300 in P3, PZ, P4, CP1 and CP2.

Splitting by Accuracy, two ANOVAs 2 (Type)*5 (electrode), one for Correct responses and other for Incorrect responses were performed. Relatively to Correct answers, effects of type $F(1,28) = 48.698$, $p < .001$, $\eta_p^2 =.635$, and electrode $F(4,112) = 6.111$, $p < .01$, $\eta_p^2 =.179$ were found, but no Interaction $F(4, 112)=1.691$, $p = .200$ was registered.

For the incorrect answers, an effect of electrode $F(4,112) = 8.275$, $p < .001$, $\eta_p^2 = .228$ and a statistical trend of type $F(1,28) = 2.986$, $p = .085$, $\eta_p^2 = .1$ were obtained. No interaction $F(4,112) = 0.899$, $p = .467$ was obtained. Relatively to a comparison of Target Correct vs Distractor Correct, significant differences were found for all electrodes (*ps* < .001). However, for incorrect answers, "Target Incorrect vs Distractor Incorrect" comparison, statistical differences in CP2 electrode $t(28) = -2.248$, $p = .033.$, 95% CI [-4.15 -2.668], Cohens'*d* =.417, Cl effect Size = .662, a statistical trend in P3 $t(28) = -2.248$, $p = .033.$, 95% CI [-4.15 -2.668], Cohens'*d* =.33, Cl effect Size=.629 were observed. No significant results were found for all the other electrodes (*ps* > .160).

Splitting by type, two ANOVA's 2 (accuracy)*5 (electrode), one for Target and another for Distractor, were performed. Relatively to Target, effects of accuracy $F(1,28) = 9.025$, $p < .001$, $\eta_p^2 = .244$ and electrode $F(1,28) = 7.494$, $p < .001$, $\eta_p^2 = .211$ were recorded, but no interactions $F(4,112) = 1.626$, $p = .210$ were found. For the distractors an effect of accuracy $F(1,28) = 25.693$, $p < .001$, $\eta_p^2 = .479$ and electrode $F(4,112) = 7.541$, $p < .001$, $\eta_p^2 = .212$ were recorded, but no interactions $F(1,28) = 0.623$, $p = .6477$ were achieved.

The comparison "target correct" vs. "distractor incorrect" showed no significant differences for P3, PZ, P4, and CP2 (*ps* > .374). Relatively to CP1 a statistical trend was observed $t(28) = 1.865$, $p = .073$, 95% CI [-.125 -2.668], Cohens'*d* = .346, Cl effect Size = .635.

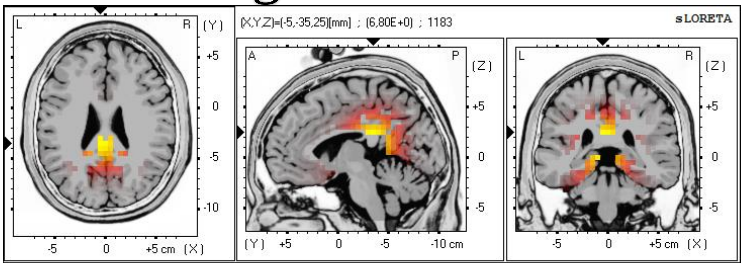
The maximum average activations in each condition in the time window of P300 are displayed in Figure 47. Before estimating the functional location of the differences between these conditions, we checked in which P300 time window latencies the *t*-test value reached a maximum. Statistical differences for "Distractor Correct vs Distractor

Incorrect” (Min t at 374 ms; $t > 3.215$, $p < .01$) and for “Target Correct - Distractor Incorrect” (Min t at 477 ms; $t > 3.688$, $p < .01$) were obtained. Log of the ratio of averages, SnPM randomizations, computed bullet proof, corrected critical thresholds and P values, and 5000 permutations were performed. Figure 48 contains a log of F-ratios for the one-tailed test: Distractor Correct < Distractor Incorrect at latency 374; and the difference of represented functional activation is significant at $p < .001$. The structures where significant differences were recorded (Voxel Value > 2.107), as well as their coordinates, are in Table 4.

Table 4. Brain structures where significant differences between Distractor Correct and Distractor Incorrect conditions were recorded – P300’s time window.

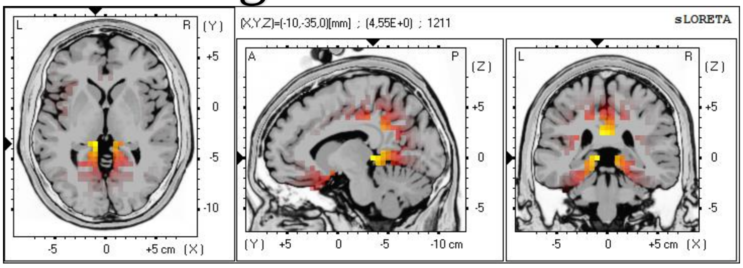
X(MNI)	Y(MNI)	Z(MNI)	X(TAL)	Y(TAL)	Z(TAL)	Voxel Value	BA	Lobe	Structure
40	20	55	40	22	50	5.39641E+0000	8	Frontal Lobe	Superior Frontal G.
45	20	50	45	22	45	5.38516E+0000	8	Frontal Lobe	Middle Frontal Gyrus
55	25	25	54	25	22	4.69229E+0000	46	Frontal Lobe	Middle Frontal Gyrus
60	10	25	59	11	22	4.67311E+0000	45	Frontal Lobe	Inferior Frontal Gyrus
55	10	25	54	11	22	4.52106E+0000	45	Frontal Lobe	Inferior Frontal Gyrus
50	25	25	50	25	22	4.49690E+0000	45	Frontal Lobe	Inferior Frontal Gyrus
-55	-35	55	-54	-31	52	4.45355E+0000	40	Parietal Lobe	Postcentral Gyrus
60	15	20	59	15	18	4.43981E+0000	44	Frontal Lobe	Inferior Frontal Gyrus
-45	-30	65	-45	-26	61	4.36849E+0000	1	Parietal Lobe	Postcentral Gyrus
-60	-35	45	-59	-32	43	4.30032E+0000	40	Parietal Lobe	Inferior Parietal Lobe
-40	-40	65	-40	-36	62	3.92236E+0000	2	Parietal Lobe	Postcentral Gyrus
-55	-45	35	-54	-42	34	3.79490E+0000	40	Parietal Lobe	Supramarginal Gyrus
50	30	20	50	30	17	3.75031E+0000	46	Frontal Lobe	Inferior Frontal Gyrus
35	5	30	35	6	27	3.45965E+0000	9	Frontal Lobe	Inferior Frontal Gyrus
-60	-20	35	-59	-18	33	3.45729E+0000	3	Parietal Lobe	Postcentral Gyrus
40	0	20	40	1	18	2.66821E+0000	13	Sub-Lobar	Insula
-65	-50	5	-64	-48	7	2.73220E+0000	21	Temporal lobe	Middle Temporal L.
50	25	0	50	24	-1	2.30081E+0000	47	Frontal Lobe	Inferior Frontal Gyrus

Target Correct



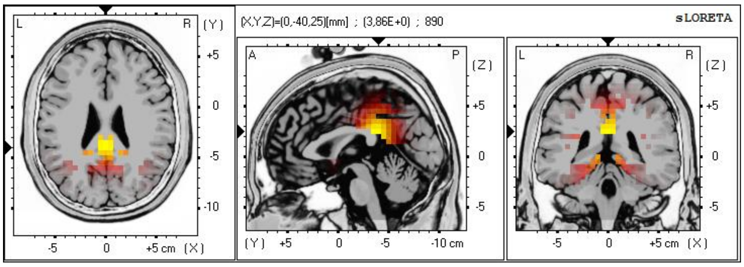
MNI			TALAIRACH						
X	Y	Z	X	Y	Z	Voxel Value	BA	Lobe	Structure
-5	-35	25	-5	-33	25	6.79952E+0000,	23	Limbic Lobe	Posterior Cing.
0	-35	25	0	-33	25	6.79170E+0000,	23	Limbic Lobe	Cingulate G.
-5	-30	25	-5	-28	24	6.77397E+0000,	23	Limbic Lobe	Posterior Cing.
-5	-40	25	-5	-38	25	6.76599E+0000,	23	Limbic Lobe	Posterior Cing.
0	-40	25	0	-38	25	6.76296E+0000,	31	Limbic Lobe	Cingulate G.
-10	-35	0	-10	-34	2	6.75738E+0000,	27	Limbic Lobe	Parahippoc. G
0	-45	20	0	-43	21	6.73459E+0000,	30	Limbic Lobe	Posterior Cing.
5	-35	25	5	-33	25	6.72237E+0000,	23	Limbic Lobe	Posterior Cing.
-10	-40	0	-10	-39	2	6.70717E+0000,	30	Limbic Lobe	Parahippoc. G
5	-40	25	5	-38	25	6.70192E+0000,	23	Limbic Lobe	Posterior Cing.

Target Incorrect



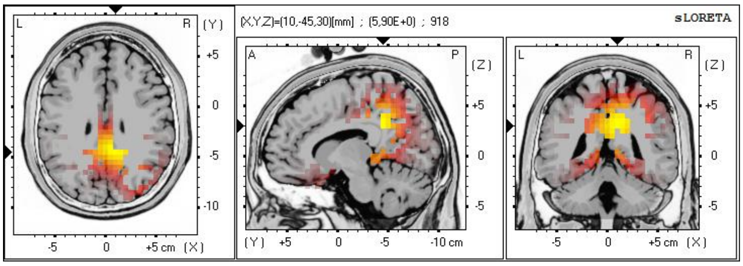
MNI			TALAIRACH						
X	Y	Z	X	Y	Z	Voxel Value	BA	Lobe	Structure
-10	-35	0	-10	-34	2	4.55391E+0000	27	Limbic Lobe	Parahippoc. G
-5	-35	25	-5	-33	25	4.55134E+0000	23	Limbic Lobe	Posterior Cing.
-5	-30	25	-5	-28	24	4.54873E+0000	23	Limbic Lobe	Posterior Cing.
0	-35	25	0	-33	25	4.53056E+0000	23	Limbic Lobe	Cingulate Gyrus
-5	-40	25	-5	-38	25	4.51662E+0000	23	Limbic Lobe	Posterior Cing.
-10	-40	0	-10	-39	2	4.50266E+0000	30	Limbic Lobe	Parahippoc. G
0	-40	25	0	-38	25	4.49552E+0000	31	Limbic Lobe	Cingulate Gyrus
-15	-35	0	-15	-34	2	4.47795E+0000	27	Limbic Lobe	Parahippoc. G
-5	-35	30	-5	-33	29	4.47361E+0000	23	Limbic Lobe	Cingulate Gyrus
5	-35	25	5	-33	25	4.47149E+0000	23	Limbic Lobe	Posterior Cing.

Distractor Correct



MNI			TALAIRACH						
X	Y	Z	X	Y	Z	Voxel Value	BA	Lobe	Structure
-5	-35	25	-5	-33	25	6.79952E+0000,	23	Limbic Lobe	Posterior Cing.
0	-35	25	0	-33	25	6.79170E+0000,	23	Limbic Lobe	Cingulate G
-5	-30	25	-5	-28	24	6.77397E+0000,	23	Limbic Lobe	Posterior Cing.
-5	-40	25	-5	-38	25	6.76599E+0000,	23	Limbic Lobe	Posterior Cing
0	-40	25	0	-38	25	6.76296E+0000,	31	Limbic Lobe	Cingulate Gyrus
-10	-35	0	-10	-34	2	6.75738E+0000,	27	Limbic Lobe	Parahippoc. G
0	-45	20	0	-43	21	6.73459E+0000,	30	Limbic Lobe	Posterior Cing.
5	-35	25	5	-33	25	6.72237E+0000,	23	Limbic Lobe	Posterior Cing.
-10	-40	0	-10	-39	2	6.70717E+0000,	30	Limbic Lobe	Parahippoc. G
5	-40	25	5	-38	25	6.70192E+0000,	23	Limbic Lobe	Posterior Cing.
-5	-35	25	-5	-33	25	6.79952E+0000,	23	Limbic Lobe	Posterior Cing.

Distractor Incorrect



MNI			TALAIRACH						
X	Y	Z	X	Y	Z	Voxel Value	BA	Lobe	Structure
10	-45	30	10	-42	30	5.89856E+0000	31	Limbic Lobe	Cingulate Gyrus
10	-45	35	10	-42	34	5.89735E+0000	31	Limbic Lobe	Cingulate Gyrus
15	-45	35	15	-42	34	5.89132E+0000	31	Limbic Lobe	Cingulate Gyrus
5	-40	30	5	-37	30	5.89040E+0000	31	Limbic Lobe	Cingulate Gyrus
5	-40	25	5	-38	25	5.88559E+0000	23	Limbic Lobe	Posterior Cing.
15	-45	30	15	-42	30	5.88363E+0000	31	Limbic Lobe	Cingulate Gyrus
5	-45	30	5	-42	30	5.88183E+0000	31	Limbic Lobe	Cingulate Gyrus
5	-45	35	5	-42	34	5.87247E+0000	31	Limbic Lobe	Cingulate Gyrus
5	-40	35	5	-37	34	5.86745E+0000	31	Limbic Lobe	Cingulate Gyrus
15	-45	40	15	-42	39	5.86333E+0000	31	Limbic Lobe	Cingulate Gyrus
10	-45	30	10	-42	30	5.89856E+0000	31	Limbic Lobe	Cingulate Gyrus

Figure 47. Top 10 activation voxels (functional localization – MNI and Talairach coord.) in P300's time window in each condition (target correct, target incorrect, distractor correct and distractor incorrect)

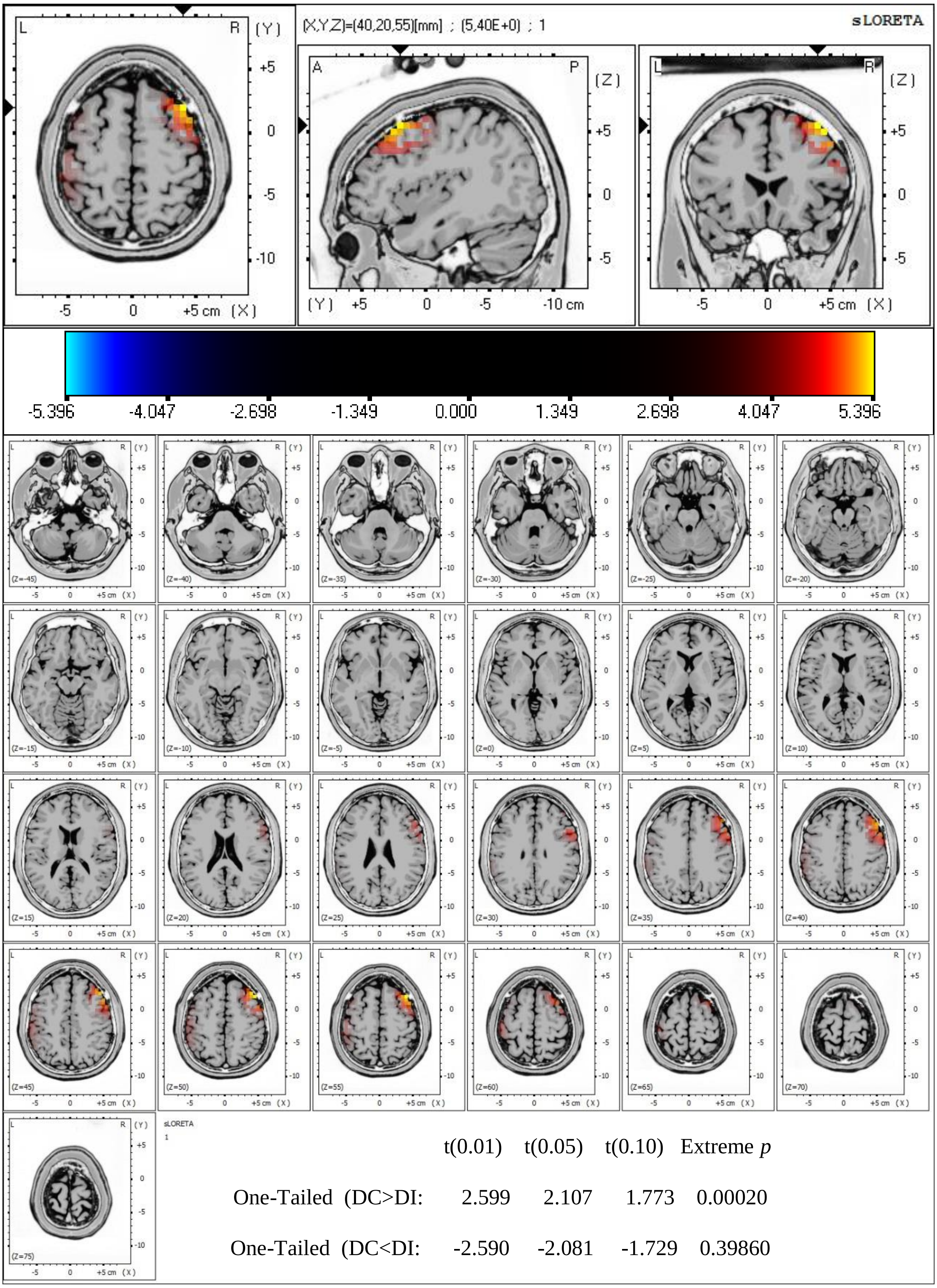
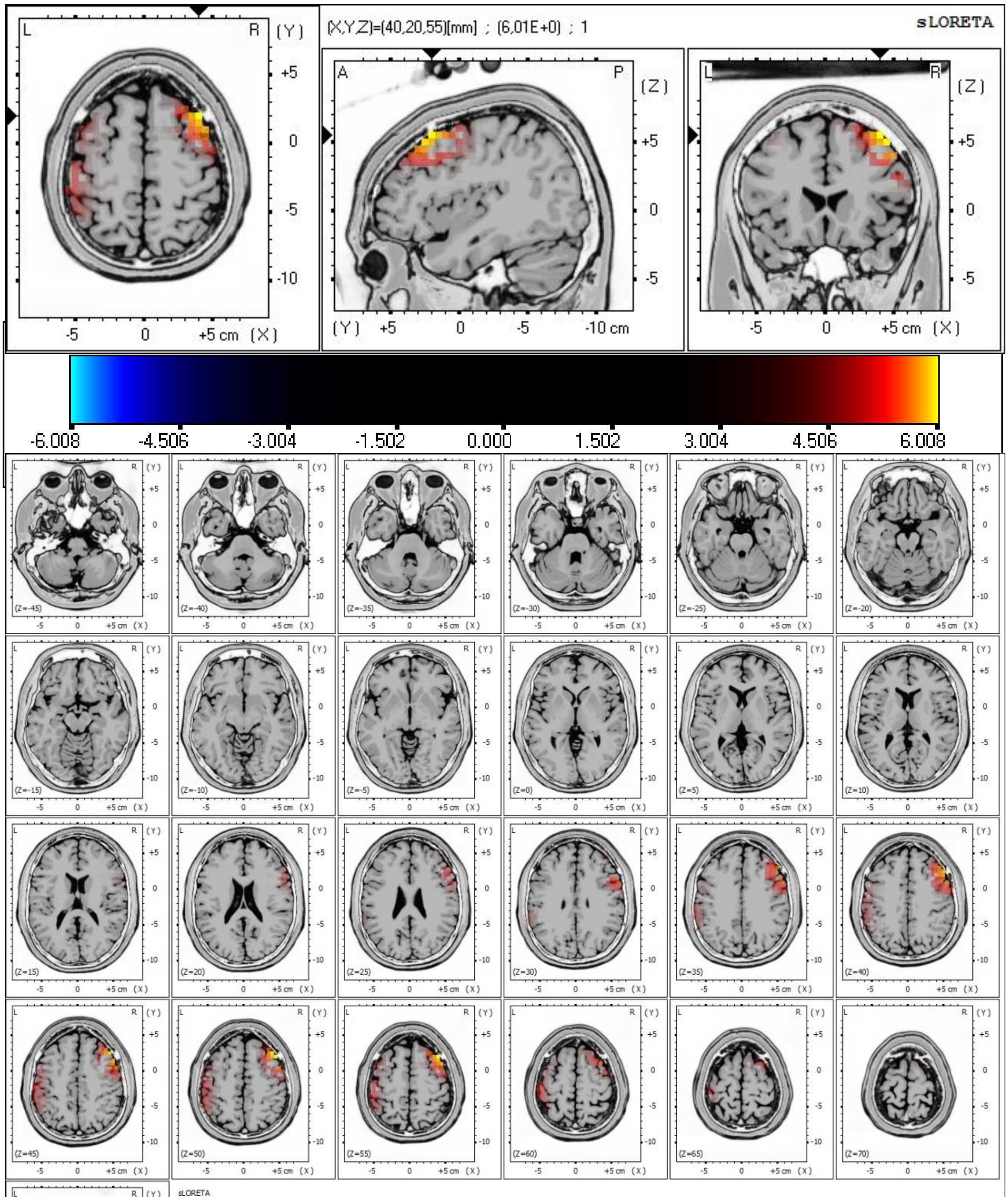


Figure 48. Log of F-ratios for the one-tailed test: Distractor Correct >Distractor Incorrect at 374 ms.

To explore the difference between Target Correct vs. Distractor Incorrect at latency 477ms the procedure recommended by Pasqual-Marqui (2012) was once again used. Figure 49 contains a log of F-ratios for the one-tailed test: Target Correct < Distractor Incorrect at latency 477; and the difference of represented functional activation is significant at $p < .001$. The structures where significant differences were recorded (Voxel Value > 1.942), as well as their coordinates, are in Table 5.

Table 5. Brain structures where significant differences between Target Correct and Distractor Incorrect conditions were recorded – P300's time window

X-MNI	Y-MNI	Z-MNI	X-TAL	Y-TAL	Z-TAL	Voxel Value	BA	Lobe	Structure
40	15	55	40	17	50	5.87111E+000	8	Frontal Lobe	Superior Frontal G.
35	20	55	35	22	50	5.81826E+000	8	Frontal Lobe	Superior Frontal G.
55	0	50	54	2	46	5.57328E+000	6	Frontal Lobe	Precentral Gyrus
55	10	35	54	11	32	5.53157E+000	9	Frontal Lobe	Inferior Frontal G.
-50	-30	60	-50	-26	57	5.29079E+000	2	Parietal Lobe	Postcentral Gyrus
55	25	25	54	25	22	5.28396E+000	46	Frontal Lobe	Middle Frontal G.
60	10	25	59	11	22	5.27865E+000	45	Frontal Lobe	Inferior Frontal G.
-50	-35	60	-50	-31	57	5.26324E+000	40	Parietal Lobe	Inferior Parietal Lob.
-55	-35	55	-54	-31	52	5.25486E+000	40	Parietal Lobe	Postcentral Gyrus
60	15	20	59	15	18	5.03469E+000	44	Frontal Lobe	Inferior Frontal G.
45	0	35	45	2	32	4.78111E+000	6	Frontal Lobe	Inferior Frontal G.
-60	-50	35	-59	-47	35	4.77672E+000	40	Parietal Lobe	Supramarginal G.
-45	-15	55	-45	-12	51	4.71898E+000	4	Frontal Lobe	Precentral Gyrus
-50	-75	15	-50	-72	17	3.05257E+000	39	Temporal Lobe	Middle Temporal G.
60	5	-15	59	4	-13	2.85756E+000	21	Temporal Lobe	Middle Temporal G.
-45	-70	30	-45	-66	31	2.66763E+000	39	Parietal Lobe	Angular Gyrus
50	15	-25	50	13	-22	2.66493E+000	38	Temporal Lobe	Superior Temp. G.
-45	-40	20	-45	-38	20	2.28264E+000	13	Sub-lobar	Insula
-40	-75	-20	-40	-74	-13	2.27614E+000	19	Occipital Lobe	Fusiform Gyrus
-50	-55	-10	-50	-54	-6	2.24260E+000	37	Temporal Lobe	Inferior Temporal G.
-50	-60	-25	-50	-59	-18	2.20804E+000	37	Temporal Lobe	Fusiform Gyrus
45	-10	-30	45	-11	-25	2.00236E+000	20	Temporal Lobe	Fusiform Gyrus



	t (0.01)	t (0.05)	t (0.10)	Extreme *p*
One-Tailed (TC>DI:	2.465	1.942	1.725	0.00020
One-Tailed (TC<DI:	-2.390	-1.944	-1.697	0.99920

Figure 49. Log of F-ratios for the one-tailed test: Target Correct >Distractor Incorrect at 477 ms.

Peripheral Measures

Heart Rate

Regarding heart rate analysis, the values of heart rate change (Heart Rate Change = heart rate $_{\text{epoch 1500ms}}$ - heart rate baseline $_{\text{500ms each item}}$) for each condition are displayed in Figure 50.

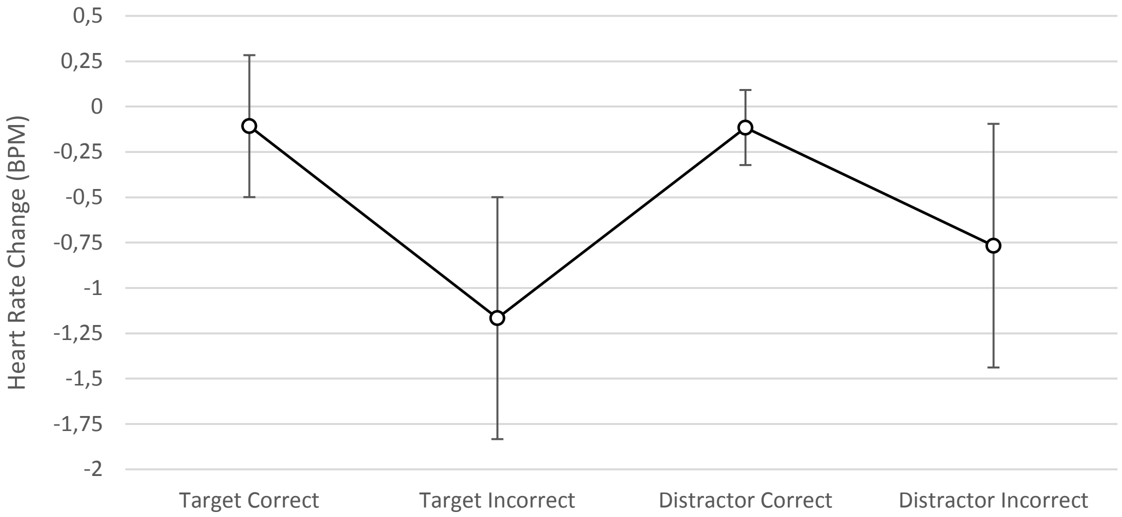


Figure 50. Heart rate change obtained in the four experimental conditions.

Although a greater deceleration for the incorrect responses was denoted, the 2 (type) * 2 (accuracy) ANOVA showed no effects of Accuracy $F(1,28) = 2.715$, $p = .111$, or type $F(1,28) = 0.233$, $p = .633$. No interaction was observed either $F(1,28) = 0.161$, $p = .691$.

Electromyography

For the analysis of the activation of the Corrugator and Zygomaticus muscles, after rectification of the signal through a mathematical integral of the biopac, the values of the area under the curve of the pre-stimulus baselines (500ms) were subtracted from the values of the area under curve recorded during the 1500 ms of the stimulus. The resulting values (change) for Zygomaticus *Major* and Corrugator *Supercili* by the condition are shown in Figure 51.

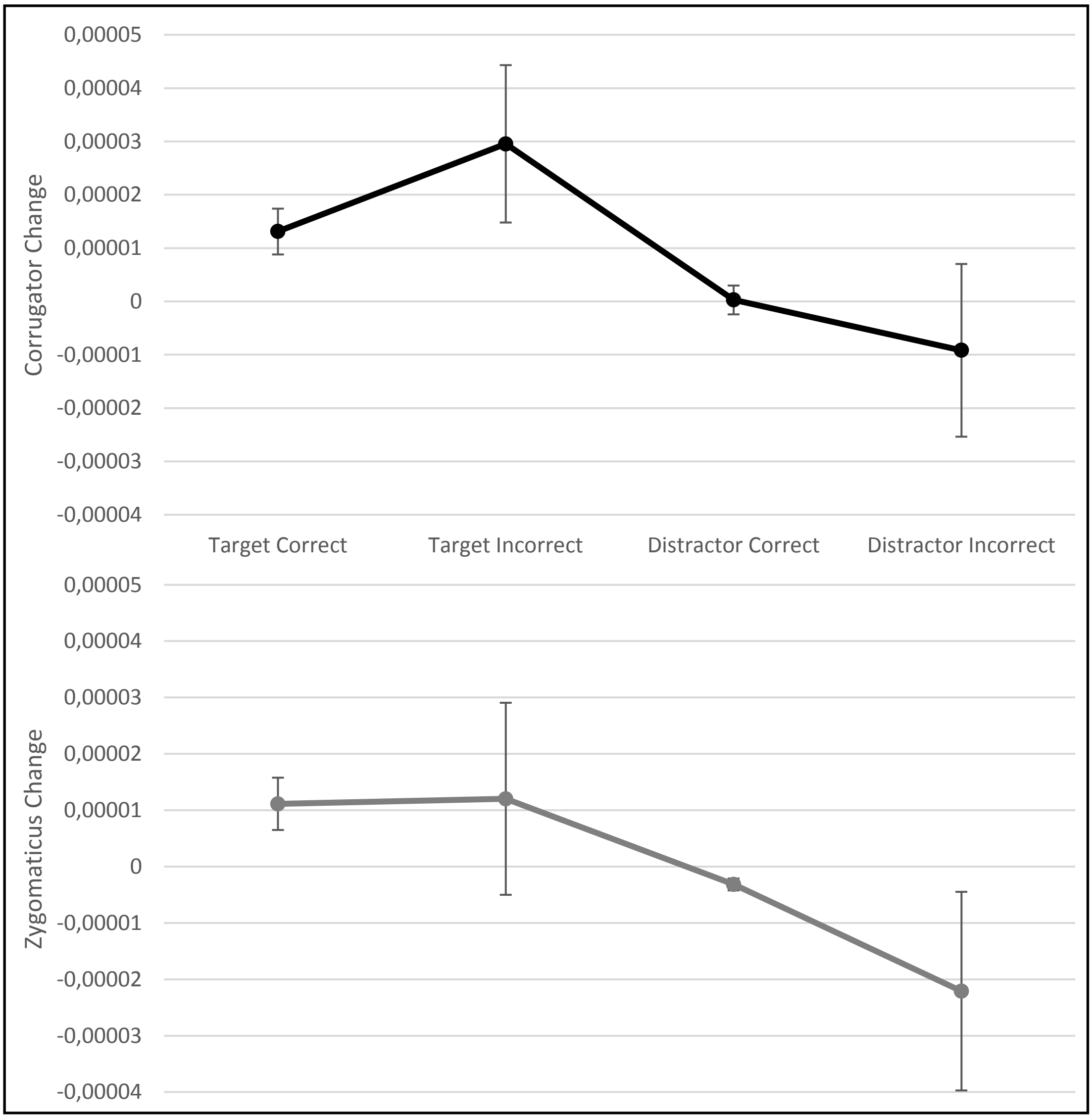


Figure 51. EMG change registered in corrugator and zygomaticus muscles in the four experimental conditions.

The Corrugator analysis shows that, regardless of the response (correct or incorrect), the targets recorded greater activation, than the distractors. Indeed, the 2 (type)* 2 (accuracy) ANOVA showed an effect of Type $F(1,28) = 10.617$, $p < .001$, $\eta_p^2 = .275$, but no effect of accuracy $F(1,28) = 0.075$, $p = .786$, or interaction of type*accuracy $F(1,28) = 2.190$, $p = .150$. Relatively to the Zygomaticus, no effects of accuracy $F(1,28) = 0.515$, $p = .479$ or interaction $F(1,28) = 0.585$, $p = .451$ were found. Generally, a similar pattern to that of the corrugator was found. However, only a statistical trend of type $F(1,28) = 3.823$, $p = .061$, $\eta_p^2 = .120$ was observed.

DISCUSSION

Eyewitnesses are often the only way to determine the identity of the perpetrator (Wells & Olson, 2003). However, the identification error rates in eyewitness studies are extremely high (Busey & Loftus, 2006; Clifford & Hollin, 1981; Coxon & Valentine, 1997) and 70% of people who were exonerated by DNA analysis were associated with cases involving incorrect identifications (Innocence Project, 2017). Due to their ethical and social consequences, the errors in eyewitness testimony are problematic (Malpass, 2006). According to Wells (2014), the use of different approaches, such as eye movements, pupil dilation or event-related potential patterns could help improve accuracy and minimize those problems. Therefore, this study aimed to explore the psychophysiological markers of eyewitness performance.

Behaviourally, the experiment has revealed that targets (the face of the perpetrator) elicited longer RTs than distractors, regardless of response accuracy. This result is in agreement with the postulated in the literature that affirms that stimuli with greater negative emotional load (unpleasant stimuli), create more interference and capture more attention, promoting longer/slower response times (e.g., Bradley, Cuthbert, & Lang, 1996; Carretié, Hinojosa, López-Martín, Albert, Tapia, & Pozo, 2009).

Concerning the P100, according to the literature, it was expected that the suspects that the participants saw committing the crime (targets), being more emotionally activating (unpleasant), would elicit greater amplitude of the P100 than the distractors – which were potentially neutral stimuli (Carritié et al., 2004; Hot et al., 2006). However, the present study found that the amplitude of the P100 was significantly larger when the participants mistakenly identified a distractor than when they correctly identified the target, or when they correctly rejected the distractor. The source estimation analysis showed that this difference in performance seems to be related to a hyperactivation in the Ventromedial Prefrontal Cortex. Although this result may be contrary to what is reported in the literature, recent studies show that the affective reaction to an emotional stimulus is altered (and may even generate an inverse pattern) when information from the amygdala is processed by the ventromedial prefrontal cortex (Buhle et al., 2014; Lapate et al., 2016). The connection of the amygdala to the prefrontal is mainly performed by a white matter trait called uncinate, and a greater connectivity in this structure has been associated with emotional regulation (e.g., Tromp et al., 2012; Westlye, Bjørnebekk, Grydeland, Fjell, & Walhovd, 2011). This emotional regulation seems not only related to the inhibitory control of the amygdalin activation to unpleasant stimuli, but to the change in emotional value, with the Ventromedial Prefrontal Cortex being a modeler of the emotional and social value attributed to a stimulus (Smith, Clithero, Boltuck, & Huettel, 2013; Winecoff et al., 2013). Adding to this explanation, Lidaka et al. 2014, in a study attempting to investigate the neuroanatomical substrates involved in true and false memories for faces, found a cluster in the medial part of the prefrontal cortex when the participants had false memories for the faces. In the same sense, Van Kesteren and Brown (2014) showed that despite the importance of the medial prefrontal cortex for memory or decision making, this structure is also related to false memories or false confidence about memories (leading to overgeneralization and over-reliance on prior knowledge). Thus, the greater amplitude of the P100 obtained when a distractor was identified to be related to a prefrontal hyperactivation, which modulates emotion and social value and is associated with false positive errors. Interestingly, the condition with the second largest amplitude of the P100 is Target Incorrect (where the difference with Distractor Incorrect showed the absence of statistical significance). This fact may again indicate that the executive modulation provided by the ventromedial is associated with the error. Future studies are needed to clarify this effect as well as the possible interaction between false positives and modulation of the amygdalin reaction.

Relatively to the N170, results show that the amplitude of this potential was significantly higher when a correct rejection was made than when a false positive occurred. Considering the result obtained by Righi et al. (2012) in which the N170 is lower (more negative) in amplitude for emotionally negative faces, this result goes against expectations. However, there appears to be a contamination of the P100 on the N170, since as in P100 the identification of a distractor causes the N170 to have a response that was expected for a target (unpleasant). In fact, due to the contamination, this difference could, again, be explained by a ventromedial hyperactivation. This result needs to be further explored in future studies, and it cannot be excluded that it can be explained by the electrical variation of the signal (greater amplitude of the P100, means less amplitude of the N170), making it necessary to calculate the variation (delta Δ) between the two components.

About the P300, a greater amplitude was expected for the target correct condition, replicating the study conducted by Lefebvre et al. (2007). In fact, the target correct condition obtained a greater amplitude in the P300 than the other conditions (target incorrect, distractor correct and distractor incorrect), but no statistical significance was reached in the target correct vs. distractor incorrect comparison. However, source analysis showed a significant difference in the time window of the P300 for this comparison. This difference was explained by a hypoactivation of the Frontal Eye

Field (FEF) when participants identified a distractor compared to when they identified the target. Remarkably, studies have shown that an inactivation of the FEF disrupts target detection during the visual search (Wardak, Ibos, Duhamel, & Olivier, 2006) and decreases visual sensitivity, which results in a greater number of false positives (Vernet, Quentin, Chanes, Mitsumasu, & Valero-Cabre, 2014).

Regarding peripheral measures, more precisely heart rate, it was expected for a deceleration to occur when participants visualized a target (criminal). Despite the greater deceleration when making a mistake (False Alarm or Miss), no statistically significant differences were found between these two conditions. Although there is literature that associates errors with a short-term deceleration of Heart Rate (e.g., Danev, Winter, 1970), another factor should be taken into account. In fact, there might have been an insufficient analysis window constrained by an epoch of 1500ms, that is, the epoch having only 1500 ms might not have been sufficiently large (as it encapsulated at most two R peaks) to capture a sustained deceleration in the heart rate. A deceleration has been recorded in larger epochs (e.g., Gamer, 2011; Hodes, Cook, & Lang, 1985), with the largest deceleration difference between neutral and unpleasant conditions found in the 3rd second after stimuli onset (e.g., Abercrombie, Chambers, Greischar, & Monticelli, 2008; Bradley, Codispoti, Cuthbert, & Lang, 2001).

Regarding electromyography, it was expected that activity over the Corrugator Supercilii would be higher in the targets than in the distractors (e.g., Allen, Harmon-Jones, & Cavender, 2001; Cacioppo, Martzke, Petty, & Tassinary, 1988). In fact, this result was verified, showing that when the participants saw a target, independently of whether or not they thought that they were guilty, it led to more activation in the Corrugator/this muscle. This fact, which is already impressive per se, by demonstrating that the autonomous nervous system, which active in response to the unpleasant stimulus, even showing a withdrawal movement, becomes more interesting when it is considered together with the central data. In this study, it seems that the peripheral nervous system is responding correctly, identifying the targets because they are emotionally more salient (negative affect), but the executive modulation provided by the prefrontal, mainly the ventromedial segment, makes errors occur. Of warning that, although not significant, the Zygomaticus, had a similar behavior to the corrugator. This "quasi-effect" might be due to the size of the epoch (1500 ms). In fact, the literature shows that Zygomatic activity in negative conditions varies greatly over 1500 ms (Mavratzakis, Herbert, & Walla, 2016), crossing the X-axis (invert signal) and showing an average activation that is higher than the neutral condition. In that study (Mavratzakis et al., 2016), the same did not happen for the Corrugator. In future studies a more detailed analysis of the epoch is necessary, breaking it into smaller segments, of 250ms (Mavratzakis et al., 2016) or 500ms (Argaud et al., 2016; Sestito et al., 2013).

Despite the importance of the prefrontal ventromedial as a socio-adaptive modeler of human behavior (eg, inhibitor of violent behaviors), it seems that, in the context of the present study, it is the interference it creates, potentially resulting from scripts of cognitive activity such as stereotypes (see Forbes & Leitner, 2014) which impairs the identification of suspects in police lineups, but provides a neuroelectric indication of a false positive. Electromyography appears to be an excellent marker for the identification of criminals. These results have to be considered carefully, and replication is needed, knowing that both their interpretation and the paradigm used, for their novelty, are conceptually and methodologically "not strong". In future studies, some considerations already mentioned above should be taken into account as well as the use of new statistical analysis packages such as multilevel analysis, to take into account all the variability of the data. Additionally, given the difficulty of syncronizing the paradigms to analyze central and peripheral measures, the central measures were prioritized in the present study, so it is necessary to isolate the peripheral measures to be studied in more detail, in particular with different time windows.

CHAPTER V

CARDIAC AND ELECTROMYOGRAPHIC PATTERNS IN EYEWITNESS TESTIMONY

INTRODUCTION

Due to the synchronization/overlapping between central and peripheral measures of the study reported in Chapter IV, certain methodological choices compromised the in-depth analysis of the latter (peripheral measures). The short epoch was the most problematic of these choices since it limited the interpretation of the peripheral measures. In fact, the 1500ms maximum epoch makes it roughly impossible to analyze measures like the heart rate due to the slower physiological response and, therefore, demand a wider time interval. Indeed, studies that measure the impact of stimuli in the heart rate often use epochs with more than 15 seconds after stimulus onset (e.g., Verschuere, Crombez, Koster, Van Bockstaele, & De Clercq, 2007; Zaitsu, 2016). Several studies, namely those that use a setting similar to the present one, use a minimum 5-second window to show a deceleration of heart rate (e.g., Colombo, 2001; Gamer et al., 2008; Richards & Casey, 1992; Verschuere et al., 2007). On the other hand, the electromyographic response is faster, and 1500ms seems to be an excessively wide interval for this kind of measure, hiding time variations (see discussion of Chapter IV). According to the literature, EMG changes occurred for faces only 250-500 ms after stimulus onset (see Mavratzakis, Herbert, & Walla, 2016).

Considering all, in the present study, which will only focus on peripheral measures, the 'epoch' for heart rate analysis will be of 10 seconds, the fixation cross (which will serve as a pre-stimulus baseline) will have the duration of 1 second, and the EMG analysis will use an epoch of 3 seconds divided into 6 intervals of 500 ms (see Argaud et al., 2016; Sestito et al., 2013). Finally, as this study can be considered a continuation of the previous one (Chapter IV), maintaining roughly the same procedure, the following Method description will only contemplate the changes that were introduced.

METHOD

Participants

Fifteen undergraduate right-handed students (8 women), aged 18 to 23 years (*M* = 20.4; *SD* = 1.32) were recruited from the University of Aveiro.

Stimuli and Task

The task was composed of 8 experimental blocks. Each experimental block (see Figure 52) consisted of a scene of a theft at the ATM, followed by a 120-second waiting period, after which the researcher provided a code that the participant typed on the computer assigned to them. Finally, five cycles of a sequential lineup with five distractor faces and a target were presented. After 5 seconds of stimulus presentation, a frame with the text “Answer Now” was showed and the participants had to decide if the face that had appeared belonged to the person who committed the crime they saw in the video, pressing “S” if it did, or "N" if it did not.

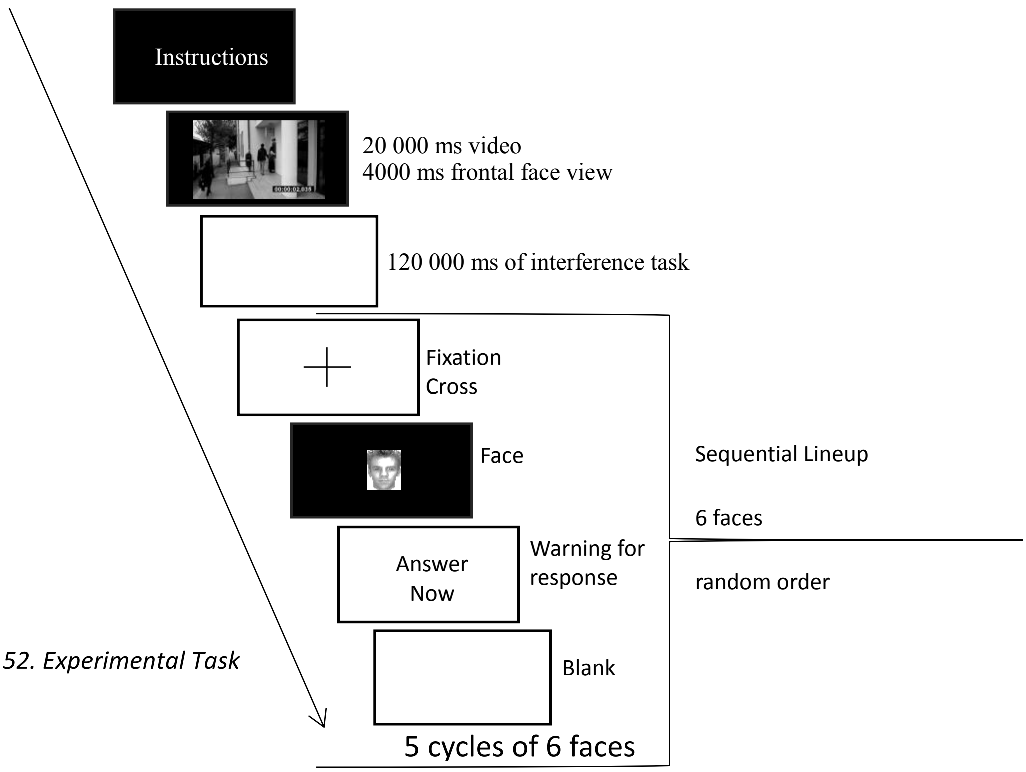


Figure 52. Experimental Task

Peripheral Analysis

With the algorithm of the Acqknowledge (Biopac) software, the heart rate (HR) was calculated from the ECG graph (QRS graph) in beats per minute. From the mean HR obtained in every second of the 10 seconds after the stimulus onset, the mean of the previous 1000 ms was subtracted (e.g., Change 1 = Average HR in Second 1 – Average HR in of the 1s baseline; Change 2 = Average HR in Second 2- Average HR in Second 1) using a heart rate change procedure (e.g., Gamer et al., 2008). EMG reactivity was measured during the 3000-ms post-stimulus period in blocks of 500 ms. Using the same paradigm (change) used for HR, to the average activity in each one of the six periods of 500 ms that together make the 3 seconds, average activity in the preceding 500 ms was subtracted (e.g., Change 1 = average of 1st block of 500ms – average of 500 prior stimuli (half baseline). The averages per measure (ECG/HR, EMGcurrutator, and EMGzygomaticus), experimental condition and participant were produced using an algorithm in Matlab (by Mathworks), purposely built for the study. Data analysis was produced with SPSS, Stata, and JASP.

RESULTS

Heart Rate

Concerning heart rate analysis, the values of heart rate change in 10 Change Blocks of 1000ms for each condition is shown in Figure 53.

Ten 2 (type)* 2 (accuracy) repeated measures ANOVAs were performed, one for each Change Block (see Figure 53, top panel, to follow this analysis). In the 1st Change block, no effects of accuracy, $F(1,14) = 1.692$, $p = .214$, type $F(1,14) = 0.253$, $p = .623$ and interaction $F(1,14) = 0.146$, $p = .709$ were observed. In the 2nd Change Block, the ANOVA showed an effect of Accuracy $F(1,14) = 5.230$, $p = .038$, $\eta_p 2 = .272$. However, no effect of type $F(1,14) = 2.193$, $p = .161$ nor interaction $F(1,14) = 3.339$, $p = .09$ were observed. Due to the visible effect in Figure 53 lower panel, found in this second for target incorrect and the meaning of a deceleration in HR reported in the literature, pairwise comparisons were performed. Statistically significant differences were found between Target Incorrect and Target Correct ($t(14) = -2.747$, $p < .01$, 95% Cl [-2.97,-.366], Cohens'd =.71, CL effect size=.76), Target Incorrect and Distractor Correct ($t(14) = -3.132$, $p < .001$, 95% Cl [-2.42, -.453], Cohens'd = .81, CL effect size = .79), but only a statistical trend between Target Incorrect and Distractor Incorrect ($t(14) = -1.863$, $p = .084$, 95% Cl[-2.86, -.201], Cohens'd = .48, CL effect size = .68) was found. Relatively to this trend, a Bayesian approach (see Figure 54) showed that although the alternative hypothesis (H1) predicts the data almost two times (1.97x) better than the null hypothesis (BF10 = 1,974, Error = 8.63e- 4) the effect is anecdotal.

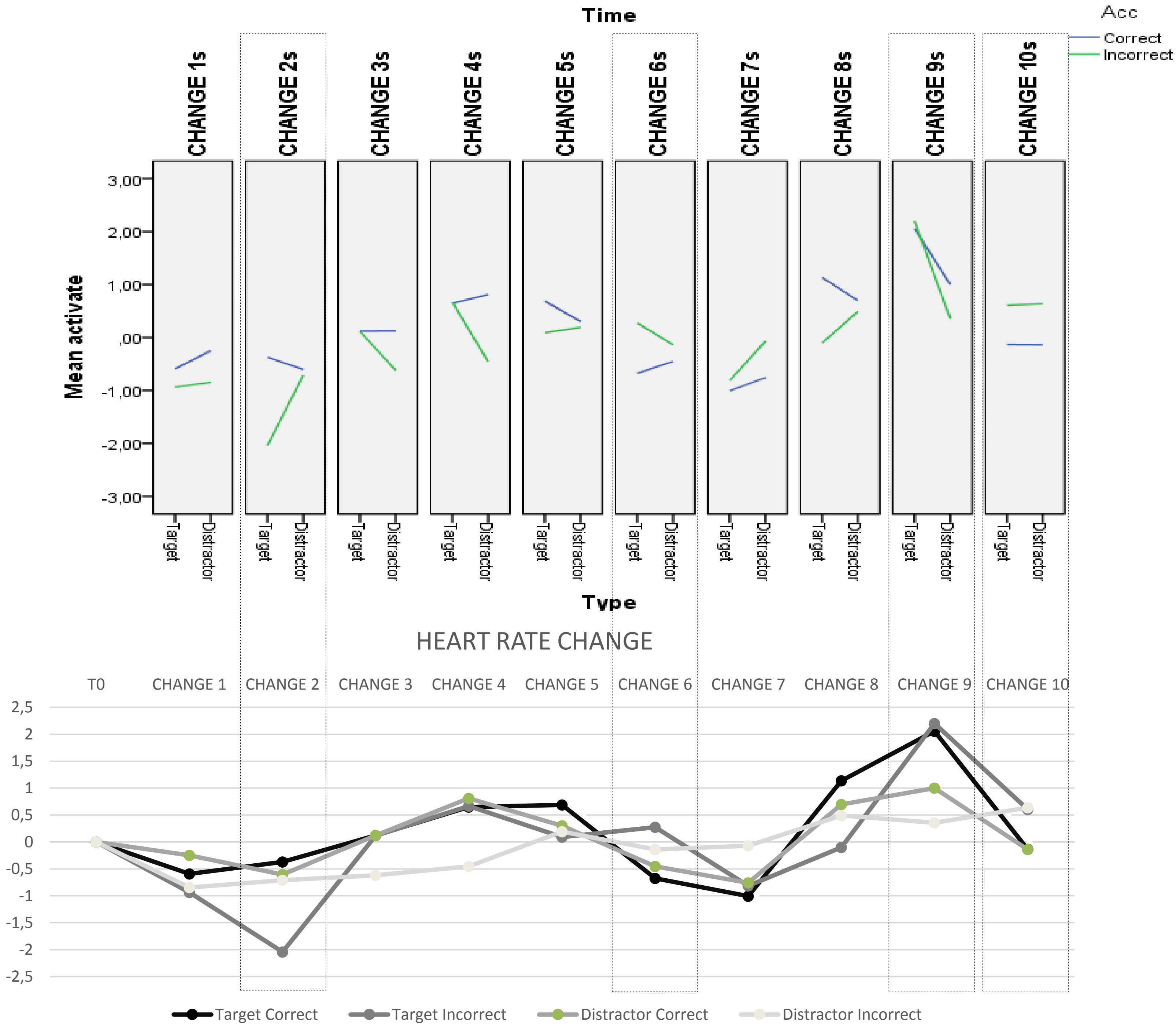
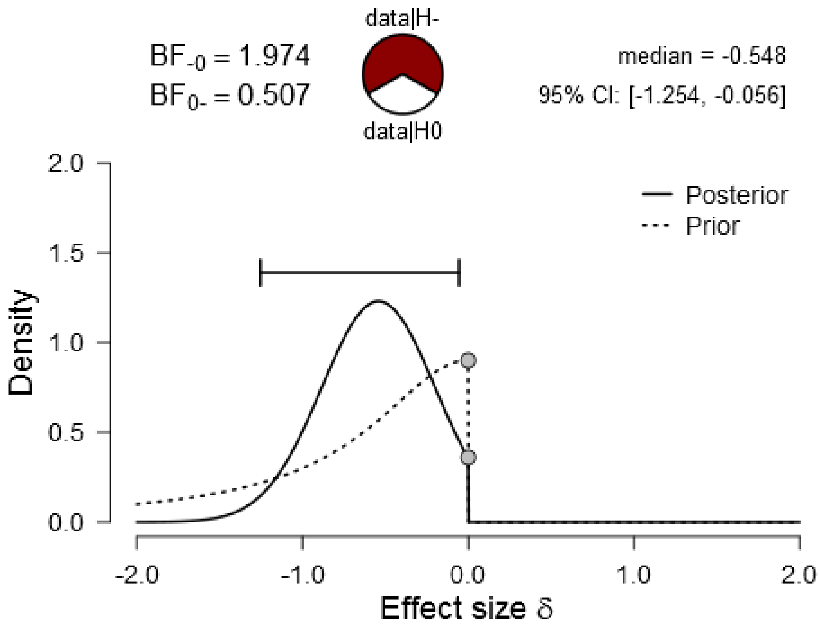


Figure 53. Heart rate change in 10 Change Blocks of 1000ms for each condition



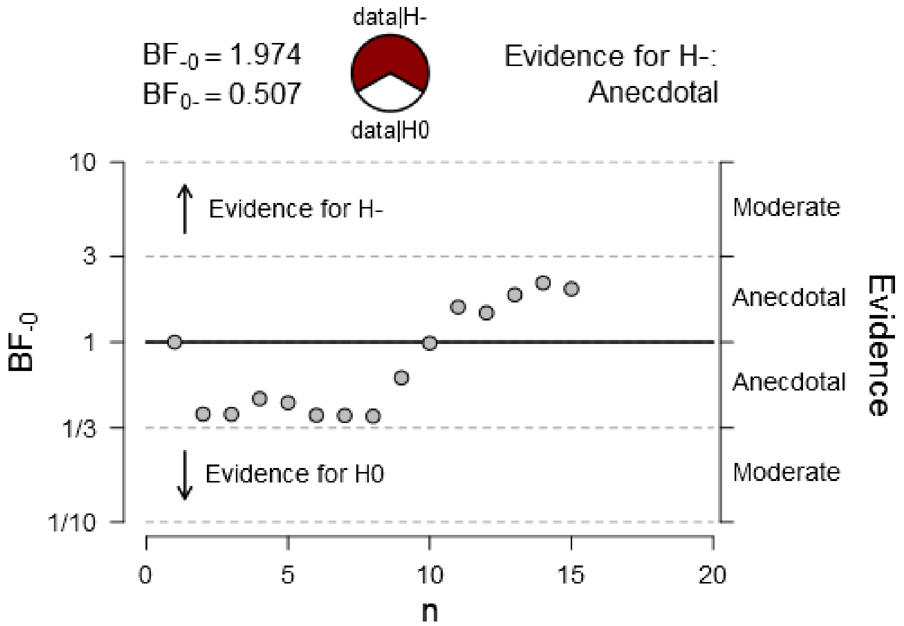


Figure 54. Top Panel - Bayesian prior distribution and Posterior distribution analysis for the relationship between Target Incorrect and Target Correct in CHANGE 2: The dot of prior distribution(red) was higher than the dot of the posterior distribution(blue) showing that the Bayes factor supports the alternative hypothesis. Bottom Panel - Sequential Analysis for the relationship between Target Incorrect and Target Correct in CHANGE 2: the supremacy of the null hypothesis was considered anecdotal.

Additionally, the HR of 10 out of 15 participants decelerated in this second in the Target Incorrect Condition (see figure 55).

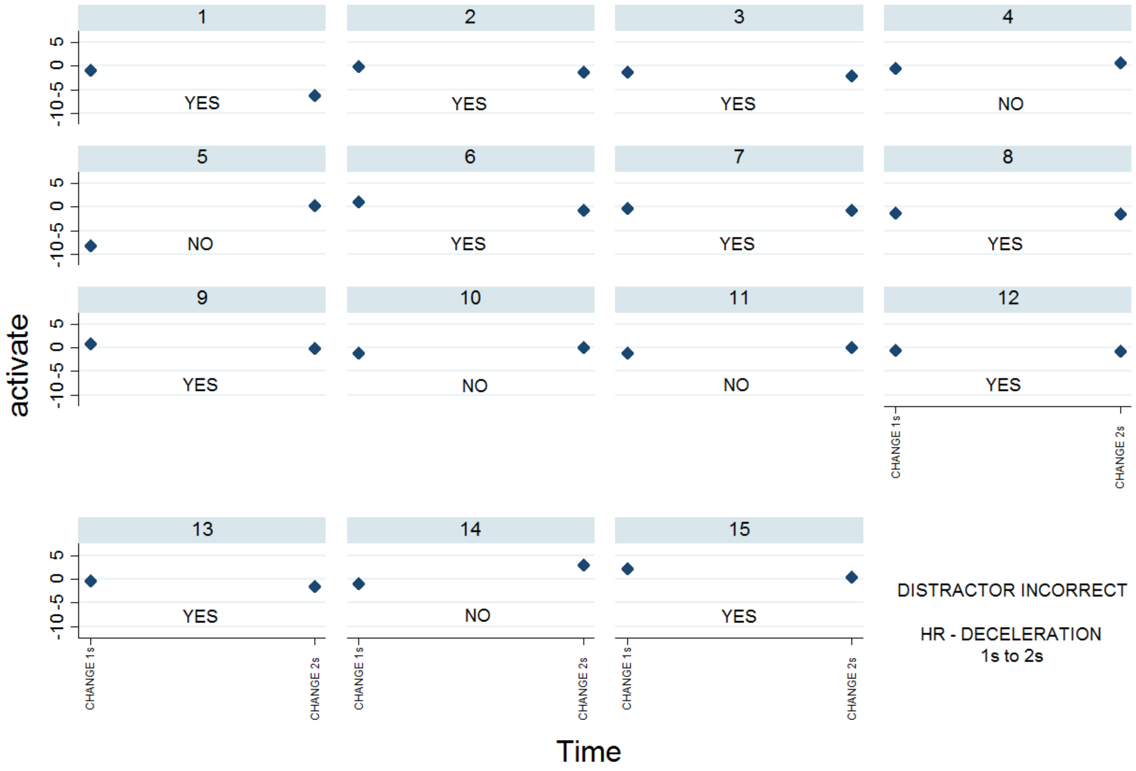


Figure 55. Heart rate deceleration in 2º second (change 2) in the Target Incorrect Condition by participant.

In the 3rd, fourth and fifth change blocks no effects of accuracy (*ps* > .158), type (*ps* > .250) or interactions (*ps* > .122) were found. Regarding the 6th Change Block, an effect of Accuracy $F(1,14) = 5.807$, $p = .030$, $\eta_p^2 = .293$ was found. However, no effect of type, $F(1,14) = .102$, $p = .754$ nor interaction $F(1,14) = .686$, $p = .422$ were observed. No effects of accuracy (*ps* > .164), type (*ps* > .163) or interactions (*ps* > .280) were found in the 7th and 8th Change Blocks. In the 9th change block, an effect of type $F(1,14) = 24.490$, $p < .001$, $\eta_p^2 = .636$, and interaction $F(1,14) = 5.991$, $p = .028$, $\eta_p^2 = .300$ were found. However, no effect of accuracy $F(1,14) = .585$, $p < .457$ was detected. The opposite pattern of statistical effects was found in 10th Change block. Indeed, in this block, an effect of accuracy $F(1,14) = 14.288$, $p < .01$, $\eta_p^2 = .506$ was verified, and no effects of type $F(1,14) = .002$, $p < .969$, nor interaction $F(1,14) = .003$, $p = .958$ were found.

In order to analyze if there is an effect of type, accuracy and time, considering all points (seconds) of the curve, a multilevel linear analysis with the following model was produced: (Intercept), Type, Accuracy, Time. The results (Corrected Quasi Likelihood under Independence Model Criterion - QICC = 1334,133) showed an effect of Accuracy, with more activity in correct answers (Wald Chi-Square = 14,389, *df* = 1, $p < .001$), an effect of time, with more activity as time increases (Wald Chi-Square = 223,174, *df* = 9, $p < .001$), and a statistical trend of Type, with more activity in the targets (Wald Chi-Square= 2,929, *df* = 1, $p = .087$).

Electromyography

The six change blocks for corrugator *supercili (*graph –A*)* and zygomaticus *major (*graph –B*)* by the condition are shown in Figure 56.

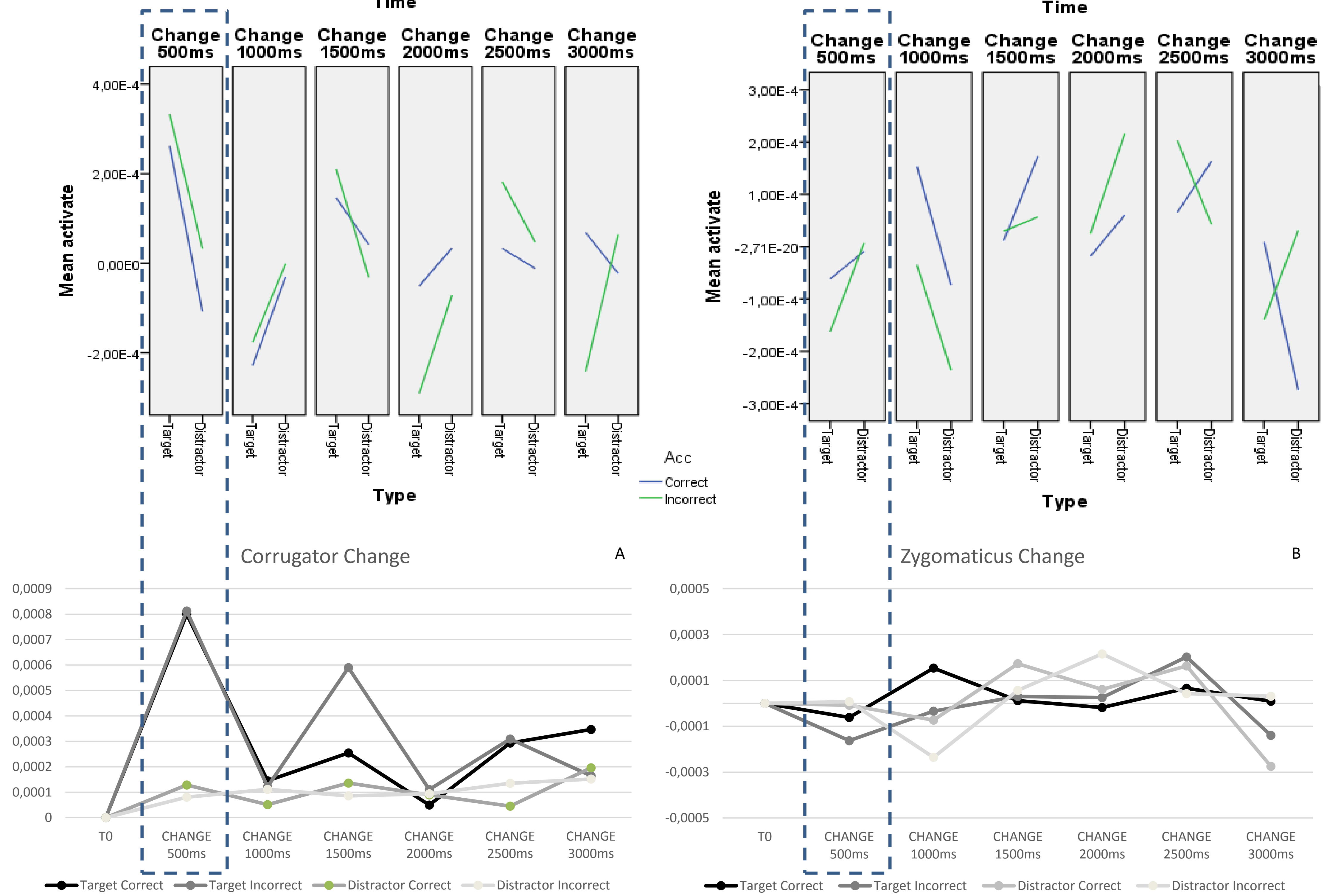


Figure 56. EMG change in 6 Change Blocks of 500ms for each condition by muscle

For each muscle, six 2 (type)* 2 (accuracy) repeated measures ANOVAs were performed.

With regard to the corrugator, for the 500 ms change block (1st Block; First 500ms – 500 of Baseline) an effect of Type $F(1,14) = 6.777$, $p = .021$, $\eta_p^2 = .326$ was obtained, but no effect of accuracy $F(1,28) = .402$, $p = .538$, or an interaction type*accuracy $F(1,28) = 0.075$, $p = .788$, were found (see graph A in Figure 56). This result indicates that regardless of the response, the activation of the corrugator for this time window was significantly higher for the targets than for the distractors.

In Change 1000 (2nd block) a statistical trend of type $F(1,14) = 3.808$, $p = .071$, $\eta_p^2 = .214$ was obtained, however no effects of accuracy $F(1,28) = 0.515$, $p = .479$ and interaction $F(1,28) = 0.585$, $p = .451$ were found. No effects (*ps* > .111) or interactions (*ps* > .141) were found for the other four change blocks.

Regarding the zygomaticus (Graph B in Figure 56) only a significant difference in the 1st (500ms) change block was achieved. Just as it happened for the Corrugator, in the 500 ms change block an effect of Type $F(1,14) = 7.988$, $p = .013$, $\eta_p^2 = .363$ was also obtained, but no effect of accuracy $F(1,14) = .982$, $p = .689$, and interaction type*accuracy $F(1,28) = 1.368$, $p = .089$, were found. This result indicates that regardless of the response, the activation of the zygomaticus for this time window was significantly higher for the distractors than for the targets. No effects (*ps* > .189) or interactions (*ps* > .137) were found for the other five change blocks.

Considering the promising results obtained for both muscles, a binary probit multilevel analysis was performed to understand the ability of a model that enters both muscles to predict the type of stimulus (target or distractor). The results (Corrected Quasi Likelihood under Independence Model Criterion - QICC = 73,348) show an interaction between muscle and activation (Wald Chi-Square = 10,050, *df* = 1, $p < .001$) with corrugator activation increasing for targets (B = 1392.157, SE = 417.49, Wald Chi-Square = 11.066, *df* = 1, $p < .001$), and zygomaticus activation increasing for distractors (B = -5276.167, SE = 2175,942, Wald Chi-Square= 5.880, *df* = 1, $p < .01$), and this model correctly classifies 78.3% of the stimuli (AUC = .783, Sensitivity = .933, Specificity = .633, see ROC curve in Figure 57).

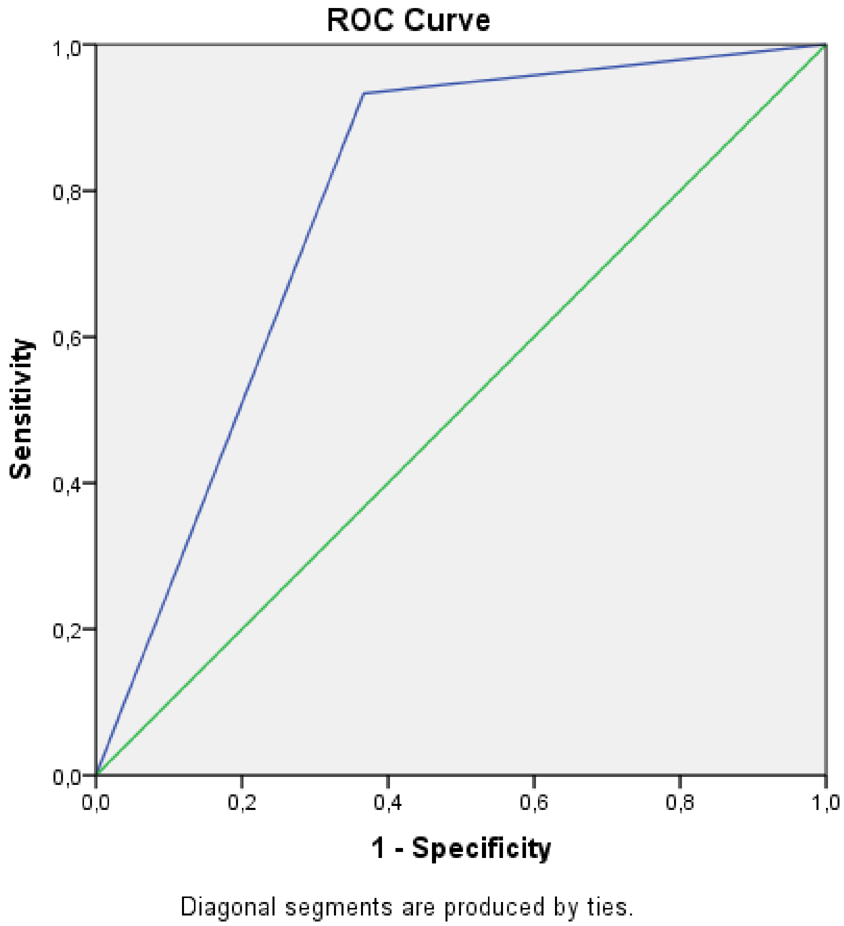


Figure 57. Blue Line: Roc Curve for Model (Corrugator + Zygomaticus); Green Line: Reference chance line

DISCUSSION

The present study aimed to verify, with a design more adequate to the in-depth analysis of peripheral measures, if the heart rate or the activation of the facial muscles corrugator and zygomaticus could be markers of performance accuracy or type of suspect (target or distractor) in an eyewitness testimony paradigm. In the same line as previous studies, a deceleration of the heart rate for the targets, for being emotionally salient, was expected (Bradley, 2009; Gamer et al., 2008). As happened in the study of Chapter IV, the results did not show a sustained deceleration towards the targets. Instead, they only showed a deceleration/slowdown in the second number 2 to the target incorrect condition. Interestingly, in the study of chapter IV, it was also the target incorrect condition that reached the greatest deceleration. This result, although not expected, can be framed within the error-related theory. In fact, there are studies that show that the errors work similarly to the negative affect in heart rate (e.g., Hajcak, McDonald, & Simons, 2004), prompting more deceleration than the correct responses (Hajcak, McDonald, & Simons, 2003, 2004). These results seem to suggest that there is a cumulative effect of the negative affect with the error and that per se the negative affect is not enough, but the error enhances the deceleration.

With regard to electromyographic activity, the literature pointed to an activity over the corrugator supercilii, greater for the targets, because they are emotionally more prominent (Larsen, Norris, & Cacciopo, 2003; Kawamoto, Nittono, Ura, 2013), and a greater zygomatic activity for the distractors, which despite not being related to positive affect, because they are neutral faces (Ribeiro et al., 2007), will promote a higher level of approach than the targets (Davidson, 1990). The observed results for the corrugator were as expected but only for the time window of 500 ms. This result is in agreement with the literature, which shows that the negative emotional faces elicited enhanced corrugator activity at 400–750 ms after stimulus onset (Künecke, Hildebrandt, Recio, Sommer, & Wilhelm, 2014; Mavratzakis et al., 2016). The same happened for the zygomaticus, that is, the expected difference was only verified in the time window of the 500 ms. Again, this result is in agreement with the study by Mavratzakis et al. (2016) which showed that at 500 ms there is a decrease in the amplitude of zygomaticus with negative emotional faces in relation to the neutral faces. When we used both muscles to try to predict which type of stimulus (target or distractor) had elicited a certain activation, the results showed that about 78.3% of the stimuli were well classified. These results seem to indicate that the electromyographic activity of the facial muscles, by responding to unpleasant emotionally salient stimuli, are good indicators that we are facing a target in an eyewitness testimony paradigm. Despite the need for further studies to explore and replicate these effects, this result is interesting, given the undesired and socially heavy consequences of eyewitness errors in this context.

CHAPTER VI

FINAL DISCUSSION AND REFLEXION

FINAL DISCUSSION AND REFLEXION

In one of the best-known forensic psychology books, Christopher Cronin, defines this area of psychology as "The application of clinical specialties to legal institutions and people who come into contact with the law" (Cronin, 2006, p. 5). It is not surprising to see this definition of forensic psychology embodied in such a well-known book since the American Psychological Association (APA) itself defines forensic psychology in the same way (see Jane Tyler Ward' article in APA site - http://www.apa.org/ed/precollege/psn/2013/09/forensic-psychology.aspx). It is, however, curious that Christopher Cronin's book cover has a full QRS graph (electrocardiography signal), but it was undoubtedly an editorial decision. Despite this lack of recognition, the importance of applying the experimental method to forensic psychology (Brigham, 1999; Dror, 2012), especially in the area of eyewitness testimony, is undeniable. (e.g.,Lindsay & Wells, 1985; Mickes, Flowe, & Wixted, 2012; Münsterberg, 1908; Wells, 2014). In fact, regarding the eyewitness domain, it all began with Hugo Münsterberg (Münsterberg, 1908) who, after graduating from Leipzig under the supervision of Wilhelm Wundt and transferring to the United States, invited by William James to manage the psychology laboratory (Münsterberg, 1922), wrote the book "On the Witness Stand", the first book on eyewitness where he raised questions about serveral topics, such as false memories, supported by experimental studies (Münsterberg, 1908) .

After Münsterberg, many important studies were developed (e.g., Bekerian & Bowers, 1983; Bregman & McAllister, 1982; Brigham, Maass, Snyder, & Spaulding, 1982; Buckhout, Alper, Chern, Silverberg, & Slomovits, 1974; Robert Buckhout, 1974; Leippe, Wells, & Ostrom, 1978; Loftus, 1979; Loftus & Palmer, 1974; Murray & Wells, 1982; Penrod, Loftus, & Winkler, 1982; Reed, 1984; Weinberg, Wadsworth, & Baron, 1983; Wells, Leippe, & Ostrom, 1979; Wells, 1984), which have allowed many improvements is procedures, such as, for example, to develop guidelines for the use of lineups (Wells, Leippe, et al., 1979). However, the quantitative and qualitative leap occurred when Roderick Lindsay and Garry Wells (Experimental Social Psychologists) introduced the experimental paradigm of absolute/relative judgment in eyewitness studies and proposed the sequential lineup (Lindsay & Wells, 1985). After this study, the supremacy of the sequential lineup has emerged, as it allows an absolute judgment, which reduces the number of **False Alarms**/**positives** by promoting the comparison of each suspect with the actual memory of the criminal. This supremacy lasted until 2012, when three known experimental psychologists of memory (Laura Mickes, Heather Flowe, & John Wixted), applying ROC curves to the analysis of the results of eyewitness studies, showed that the sequential procedure appears to be inferior to the simultaneous procedure in discriminating between the presence *versus* absence of a guilty suspect in a lineup (Mickes et al., 2012). In the following years, these results were replicated in other studies using the same methodology (ROC curves; e.g., Gronlund et al., 2012; Wixted & Mickes, 2015). Although the simultaneous lineup promoted a greater number of **Hits**, the sequential lineup continued to be known as the one responsible for the lower number of false positives (e.g., Wells, 2014). These contradictory results created a deadlock in the literature that led Garry Wells, a prominent author in eyewitness research, to ask "Should we adopt a new procedure (simultaneous) that increases the chances that the guilty might be identified but also increases the chances of mistaken identification?" (2014, p. 14). Wells, in his article in the journal Current Directions in

Psychological Science, in 2014, argued that the answer to this and other questions in eyewitness, such as "how to reduce the number of errors in identification" seems to be in the new methods used in experimental psychology, affirming: “a wholly different approach probably would be developed, perhaps one involving eye movements, pupil dilation, event-related potential patterns, response latencies, implicit memory tests, and other potential indicia of recognition. Bringing psychological science to bear on the serious problem of eyewitness identification ought to mean much more than manipulating whether photos are shown as groups versus one at a time. The next generation of eyewitness researchers should throw out the traditional” (p. 15).

Heather Flowe (the same experimental psychologist who was involved in the team that made use of the ROC curves to analyse the effectiveness of a lineup) has made significant attempts at the use of these new methodologies, more precisely analysing the ocular patterns (Flowe, 2011; Flowe & Cottrell, 2011). However, the results were not encouraging, showing ocular patterns as a weak predictor of a correct eyewitness identification. Regarding the psychophysiological patterns of accuracy in eyewitness testimony, as far as we are aware, only one study was conducted. In this study, Christine Lefebvre (Lefebvre et al., 2007), showed that the amplitude of the P300 (evoked potential that appears between 300 and 600 ms after the onset stimulus), was significantly higher when participants visualized a target /criminal than when they saw a distractor.

The present Ph.D. work was motivated by the research needs that were inherent to the alert call made by Wells (2014) and by the scarcity of literature in this specific field. Thus, the work reported in the present Ph.D. thesis was conducted in an experimental forensic framework (I think that the APA will not fine me for this reverie!), using eyetracking, psychophysiological responses, and response patterns and latencies to answer two important questions in the field of eye testimony. The first question was centred on the "Sequential vs. Simultaneous" discussion and questioned whether the merging of the two types of lineup, with a simultaneous lineup performed after the sequential lineup with stopping rule, would allow combining the best characteristics of the two lineups. That is to say, we tried to investigate whether the absolute judgment of the sequential lineup with stopping rule (responsible for the lower number of false positives) would be maintained in the simultaneous lineup and whether this lineup (simultaneous), with its relative judgment, increased the number of **Hits** (i.e., correct identification of the targets). Thus, the existence of a mixed judgment (the absolute judgement, carried over from the previous sequential lineup, and relative judgement) in the simultaneous lineup was tested. Secondly, we questioned whether there were any psychophysiological markers of correct and incorrect judgments or the visualization of a target in eyewitness testimony.

Regarding the first question, the results of the study reported in Chapter III showed that a mixed procedure in two phases, consisting of a simultaneous lineup (2nd phase) after a sequential lineup with stopping rule (1st phase), seems to let you combine the best of both types of lineup. In fact, the recollection process (which reflects the retrieval of qualitative information about a specific study episode) used in the sequential lineup for allowing the absolute comparison of the seen suspect with the existing memory of the criminal (absolute judgment), contributes to a lower number of false positives (Flowe & Ebbesen, 2007; Gronlund, 2005), and does not seem to have been entirely replaced by a familiarity processing in the simultaneous lineup (second phase), but by a somewhat mixed processing. In the second phase, there seemed to be a maintenance of the recollection, but the increased sense of familiarity (particularly considering that the stimuli are present at the same time; Meissner, Tredoux, Parker, & MacLin, 2005), consistent with the relative judgment decision strategy (Lindsay & Wells, 1985), has made the participants more successful, achieving a greater number of **Hits** than when the simultaneous lineup was presented alone. This effect was supported not only behaviourally, but also at the level of eyetracking measures. Behaviorally, the effect was supported by maintaining the low number of false positives from the first to the second phase (mostly in the target absent condition), and by increasing the number of **Hits** on the target in the second phase. At the level of the eye tracking variables, the effect was based on the maintenance of the high values of dwell time on the faces of the lineup (absolute judgment indicator - Flowe, 2011) from the first to the second phase in the categories **Misses** (despite a trend) and **False Alarms** in the target absent condition, those categories that do not behaviorally change from the first to the second phase. The dwell time pattern in the target and distractor in these categories (**Misses** and **False Alarms** - more dwell time in the distractors than in the target) was also maintained from the first to the second phase. Interestingly, this was only verified for the simultaneous lineup after the sequential stopping rule, not being verified in the simultaneous lineup without any lineup before, nor in the simultaneous lineup with the passive sequential lineup. Although promising, the results related to the fusion of the lineups will have to be explored in further studies and replicated. Nevertheless, in our opinion, they are a valuable increment to the literature on eyewitness testimony.

Regarding the second question (Are there any psychophysiological markers of accuracy, or of the visualization of a target or even a false alarm/positive in eyewitness testimony?), we expected a greater activation in the P100 (Bar-Haim et al., 2007; Carretié, 2014; Carretié et al., 2004), in the P300 (Olofsson & Polich, 2007; Righi et al., 2012), and in the corrugator (Cacioppo et al., 1986; Larsen, Norris, & Cacioppo, 2003). We also expected a deceleration in heart rate (Bradley, 2009; Gamer et al., 2008) and an amplitude decrease in zygomaticus activity (Mavratzakis et al., 2016) for targets (criminals) because they are emotionally salient, unpleasant and elicit negative affect and motivation to withdrawal. The results showed some inconsistencies in relation to what was expected. Regarding EMG recordings, activity in the muscles corrugator and zygomaticus revealed the expected pattern, suggesting that it might be possible to consider them as a somatic markers of target visualization (for the corrugator, the result of the study of Chapter IV was replicated in the study of Chapter V). In fact, corroborating the literature (e.g., Larsen et al., 2003; Mavratzakis et al., 2016), a greater amplitude activation in the corrugator and a smaller amplitude of the zygomatic activity were registered within the first 500 ms, when the participants visualized a target, independently of whether they identified it as the criminal or not (hit or miss).

According to the literature, it was also expected that the individuals who the participants saw committing the crime (targets), because they were more emotionally activating (unpleasant), would elicit greater amplitude of the P100 than the distractors - potentially neutral stimuli (Carritié et al., 2004, Hot et al., 2006). However, contrary to what was expected, the P100 proved to be a good marker of false positives, with a greater activation elicited by the distractors when they were incorrectly identified as the target, and this activation was concomitant with a hyperactivation in ventromedial prefrontal areas. Although this result may be contrary to what has been reported in the literature, recent studies show that the affective reaction to an emotional stimulus is altered (and may even generate an inverse pattern) when information from the amygdala is processed by the ventromedial prefrontal cortex (Buhle et al., 2014; Lapate et al., 2016). In fact, it seems that there was a change in emotional value, with the ventromedial prefrontal cortex being a modeler of the emotional and social value attributed to a stimulus (Smith et al., 2013; Winecoff et al., 2013). Within another framework (face recognition) but still interesting, Lidaka et al. 2014, found a cluster in the medial part of the prefrontal cortex when the participants have false memories for the faces.

For the P300, a greater amplitude for the targets than for the distractors was expected (Lefebvre, Marchand, Smith, & Connolly, 2007; Righi et al., 2012). Although there was a greater activation on the Target Correct condition, this was very close to the activation on the Distractor Incorrect condition (second largest and no statistical difference with the first). This result suggested that, unlike the P100, in the time window of the P300, there is already a greater "recognition/attention to" of the displeasure provoked by the target, although a great activation in the condition of false positives continues to occur. Despite the non-existent difference between Target Correct and Distractor Incorrect conditions with respect to the P300, in their time window, functional differences were found in the Frontal Eye Field (FEF), with a hypoactivation registered in the distractor incorrect condition. In fact, FEF hypoactivation seems to reduce the capacity of item processing (Vernet, Quentin, Chanes, Mitsumasu, & Valero-Cabre, 2014) and the ability to detect targets (Wardak et al., 2006).

Regarding heart rate, a "fear bradycardia" was expected, that is, a deceleration for emotionally relevant, threatening or unpleasant stimuli (Bradley, 2009; Gamer et al., 2008). However, there was only a significant deceleration in the Target Incorrect condition (the target was not seen but identified as such), showing the need to accumulate the effect of the error (Hajcak et al., 2003, 2004) with the effect of the unpleasant stimulus (target) to decelerate. Considering that the heart rate is the result of sympathetic and parasympathetic activation (sympathetic-parasympathetic balance; Ekman, Levenson, & Friesen, 1983; Friedman & Thayer, 1998; Quintana, Guastella, Outhred, Hickie, & Kemp, 2012), and that the worst performance in a cognitive task (error) is related to a higher preponderance of the sympathetic nervous system (Luque-Casado, Perales, Cárdenas, & Sanabria, 2016), one could argue that it is the sympathetic nervous system, which is related to arousal (e.g., Bradley, Miccoli, Escrig, & Lang, 2008) that is producing the deceleration (as unlikely as this might seem). Thus, without the existence of the error, the negative valence and the arousal of the target do not seem sufficient to produce the expected deceleration. In fact, these results seem to contradict the literature, since cardiac acceleration is associated with sympathetic activation and deceleration to parasympathetic activation (e.g., Berntson, Boysen, Bauer, & Torello, 1989; Berntson, Quigley, Lozano, Cacioppo, & Tassinary, 2007; all medical literature). Even so, it also seems paradoxical that individuals with elevated tonic heart rate, reflecting a heightened state of arousal, are more likely to show greater heart rate deceleration during processing of emotional stimuli (Abercrombie et al., 2008). Also, individuals who feel threatened, inhibit the vagal response, triggering the set of responses to survive (mostly dependent on sympathetic activity), and when the environment is safe, the "vagal brake" is applied, encouraging homeostatic functions (Quintana et al., 2012). In addition, recent research, using heart rate variability, showed that with low temporal scales (during the first second; Pan et al., 2016), both

Deceleration Capacity (DC) and Acceleration Capacity (AC) were solely dependent on vagal activity. However, with high temporal scales (time window = 3seconds; Pan et al., 2016; Ursino & Magosso, 2003), both DC and AC were positively correlated with sympathetic activity. Interestingly, the Porges model (Porges, 1992) argued that physiological response to a stimulus is split into two components, reactive and sustained. The reactive component is vagally-mediated and occurs within the first second of stimulus perception. The sustained component, which can last 2 to 5 seconds, is mediated by a coupling of vagal withdrawal and sympathetic activation. The initial deceleration observed in the present study when an incorrect response is given to a target (i.e., a miss occurs) is thus hard to explain within the current theoretical framework.

Although the study of Chapter III was constructed to answer the first question more directly, it also showed us that, for the sequential lineup, the dwell time on the target was significantly greater when the participants identified it than when they did not identify anyone (**Misses**), or when they identified an innocent (**False Alarms**). This result can be explained by the fact that in the **Hits** the stronger memory trace (Steblay et al., 2001), and a feeling of recollection (Flowe & Ebbesen, 2007; Gronlund, 2004) promote more dwell time on the target's face. Relatively to the simultaneous lineup, when participants identified the target (**Hits**), they made significantly more visits to the target than to the average of the distractors. This result can be explained by the type of judgment used in the simultaneous lineup, i.e., when a participant is facing a set of faces that appear simultaneously and selects a face as the potential target, the relative judgment causes him/her to enter an elimination process, comparing the other faces with the selected (Dunning & Stern, 1994 – target->distractor 1->target->distractor 2-

>target ... until decision). Thus, the selected face receives in total a higher number of fixations/visits than the average of the distractors. Although salient stimuli promote longer dwell times and more visits (Simola, Torniainen, Moisala, Kivikangas, & Krause, 2013), unfortunately, these data cannot be integrated into the discussion with other chapters (IV e V) due to the use of distinct lineups/paradigms (sequential lineup with stopping rule and simultaneous).

Proposed explanatory model[24] (see figure 58)

Based on the global results of the present work and considering the related literature, we propose a provisional explanatory model, that attempts to congregate all the data that was collected in the various studies.

It is possible that when we see a target in a lineup, because it is emotionally salient, and regarded as possibly threatening, it requires a rapid detection and visually guided response, being that information sent to the amygdala by the Superior Colliculus-Pulvinar Pathway – fast track, instead of using the long track (Lateral Geniculate –Visual Areas– Inferior temporal Areas pathway) typical pathway in visual processing (Soares, Maior, Isbell, Tomaz, & Nishijo, 2017). After the information arrives at the amygdala, a hyperactivation occurs, which triggers a fight or flight reaction (e.g., Pessoa & Adolphs, 2010), that follows by a white matter trait called uncinate, to the ventromedial prefrontal cortex, (e.g., Tromp et al., 2012; Westlye, Bjørnebekk, Grydeland, Fjell, & Walhovd, 2011). Upon reaching the ventromedial prefrontal cortex from the amygdala,

[24] In this model only the arrows that are directly related to the studies carried out in this thesis are represented.

the information is modeled by modifying the colouring of the emotional response (e.g., Buhle et al., 2014), making the targets be perceived as non-salient stimuli (see Lapate et al., 2016). This ventromedial modulating mechanism (see Smith, Clithero, Boltuck, & Huettel, 2013; Winecoff et al., 2013) is related, in the present study, to the exacerbation of the emotional response to a distractor (false positive) resulting from a hyperactivation in this area (Chapter IV). Interestingly, this hyperactivation of the ventromedial prefrontal cortex was related to a later hypoactivation (+-400ms) of the frontal eye field (Chapter IV) which usually translates into a reduced ability to detect targets (Wardak et al., 2006). The frontal eye field appears in the model of Figure 58 connected to the visual areas V4 and V3, as suggested by Pessoa & Adolphs, 2010's model, and because they have a later impact on the results. However, studies using diffusion tensor imaging (DTI) present evidence that this area is also in close connection to the Superior Colliculus-Pulvinar Pathway (Tamietto, Pullens, De Gelder, Weiskrantz, & Goebel, 2012).

In addition to the information sent to the ventromedial prefrontal cortex, the amygdala firing due to target visualization is also sent by projections that leave the basal nucleus of the amygdala to the supplementary motor area (Lang, Bradley, & Cuthbert, 1998) and anterior and caudal midcingulate cortex area (Picard & Strick, 2001), which controls the muscles of the upper face (Gothard, 2014), and to the ventrolateral premotor cortex (Mushiake, Saito, Sakamoto, Itoyama, & Tanji, 2006) which, in conjunction with projections of the caudal area of the midcingulate cortex, controls the muscles of the lower face (Gothard, 2014). When the fight or flight information arrives in upper and lower half facial muscles, there is a hyperactivation in the corrugator and a hypoactivation in the zygomaticus (Chapter IV and V). Although this hyperactivation and hypoactivation of the facial muscles results from the functioning of the somatic nervous system (Cacioppo, Tassinary, & Berntson, 2007), there seems to be, considering its speed, an autonomous component in this response, mediated, for example, by the neural network anterior midcingulate cortex – autonomic system (see Gothard, 2014). In fact, findings from patients that suffered strokes suggest the existence of an alternative "limbic" pathway that controls emotional facial expressions. Patients with strokes with commitment in primary motor and premotor areas cannot produce voluntary smile when asked to do so, but can smile in response to a joke (e.g., Dawson, Hourihan, Wiles, & Chawla, 1994; Trepel, Weller, Dichgans, & Petersen, 1996), while patients with the midcingulate area affected, can produce voluntary facial movements but they are not able to produce spontaneous emotional expressions (e.g., Feiling, 1927).

The amygdala firing also follows to the autonomous nervous system by the hypothalamus–nucleus tractus solitarius pathway (Berntson et al., 2007; Rogers, Rybak, & Schwaber, 2000) that, by projecting information to the dorsal motor nucleus of the vagus nerve and to the nucleus ambiguus, controls the parasympathetic nervous system (Jordan, 2005), and by the tract of the median raphe nucleus, that together with the rostral ventrolateral medulla controls the sympathetic nervous system (Agorastos, Kellner, Baker, & Stiedl, 2014; Underwood, Arango, Bakalian, Ruggiero, & Mann, 1999). Even though this firing reaches both the sympathetic and parasympathetic nervous systems, when an absence of identification occurs in the presence of the target, we observed a sustained heart rate decelaration (Chapter V). Based on the reviewed literature, the erroneous response is likely to trigger a sympathetic response, although the vizualization of the potentially threatening target stimulus was expected to trigger a parasympathetic response. Given that no decelaration occurred when a correct response was given to the target (hit), one could assume that the observed deceleration was due to the sympathetic predominance in cardiac activity in response to the error. This seems, however, paradoxical to all the physiological and medical literature, and warrants further investigation.

Thus, by bringing all the methods above and its markers together, the model, despite new, seems to provide a relevant contribution to help detecting eyewitness identification errors and its catastrophic social consequences.

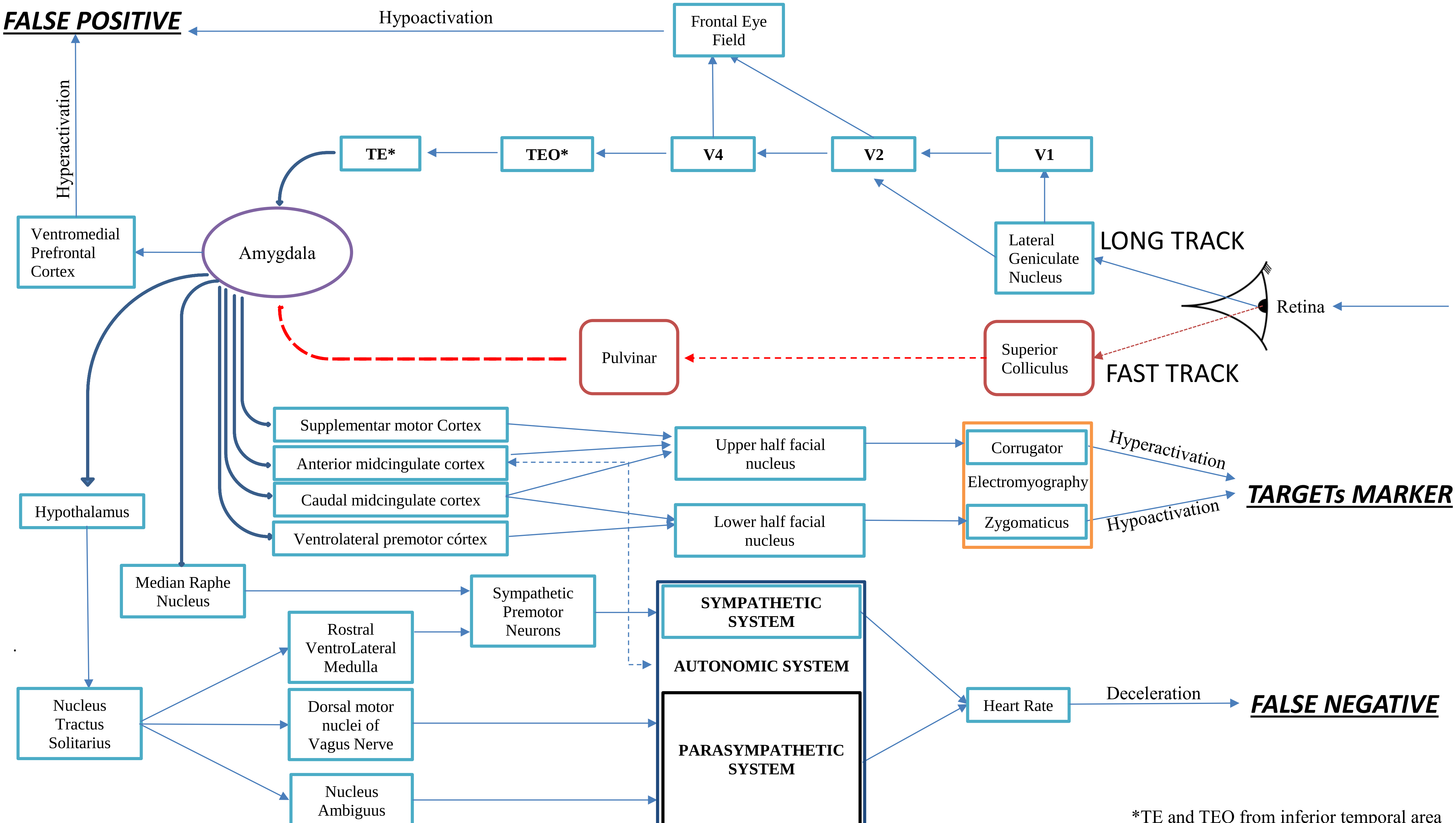


Figure 58. Proposed explanatory model

Some of the limitations of the present investigation have already been mentioned in detail in the corresponding chapters. The most prominent limitation regards the eyetracking study in which, by using the average of all the distractors to compare with targets, we may have lost important information about the particular stimuli that commanded a bigger influence on the results. That is, if one particular 'distractor' face had a significantly higher dwell time than all the others, we could be unaware of potential variables that are modulating this phenomenon and, hence, be unable to speculate any further. Another constraint of the present work is related to the fact that, even though stimuli were previously rated on attractiveness and distinctiveness, they were not rated on arousal or valence, which are commonly evaluated in emotional processing studies. Furthermore, like in many other investigations, there was a relatively short number of participants involved, which, despite reducing the probability of a type-I error, did not allow for the optimal power of statistical tests.

Future studies will have to be performed to try to replicate the results obtained, as well as to test the proposed explanatory model (see Figure 58), using an experimental paradigm and studies of central, peripheral and ocular measurements. An example of a future study of vital interest would aim to test if, by limiting awareness (for example, through continuous flash suppression - CFS), the event-related potential P100 would stop signaling false positives. Additionally, after a conceptual and methodological consolidation in an experimental setting, the model could go through an ecological validation, in order to, later, be applied in the legal context, when eyewitness evidence is collected. Additionally, new and innovative techniques might start to be applied to these research paradigms, such as the use of wireless devices to record psychophysiological measures that are even less intrusive to the participant and potentially more valid from the ecological standpoint. This is the case of bioradars, which allow to record heart rate and respiration cycles at a distance and wirelessly. Such a setup is illustrated in Figure 59, from a pilot experiment using an eyewitness paradigm that has been carried out in our lab, with very promising preliminary results. Of course that there is still a long way to go, but as is a common saying in Portugal “o caminho faz-se caminhando” (“the path is made by walking it”).

Figure 59. Bioradar setup from a pilot study with an eyewitness paradigm.

CHAPTER VII

REFERENCES

Abercrombie, H. C., Chambers, A. S., Greischar, L., & Monticelli, R. M. (2008). Orienting, emotion, and memory: Phasic and tonic variation in heart rate predicts memory for emotional pictures in men. *Neurobiology of Learning and Memory*, *90*(4), 644–650. doi: 10.1016/j.nlm.2008.08.001

Adams, R. B., Gordon, H. L., Baird, A. A., Ambady, N., & Kleck, R. E. (2003). Effects of gaze on amygdala sensitivity to anger and fear faces. *Science*, *300*(5625), 1536. doi: 10.1126/science.1082244

Adolphs, R. (2008). Fear, faces, and the human amygdala. *Current Opinion in Neurobiology*, *18*(2), 166–172. doi: 10.1016/j.conb.2008.06.006

Agencies, L. E. (2013). A National Survey of Eyewitness Identification Procedures in Law Enforcement Agencies.

Agorastos, A., Kellner, M., Baker, D. G., & Stiedl, O. (2014). Diminished Vagal and/or Increased Sympathetic Activity in Post-Traumatic Stress Disorder BT - Comprehensive Guide to Post-Traumatic Stress Disorder. In C. R. Martin, V. R. Preedy, & V. B. Patel (Eds.) (pp. 1–15). Cham: Springer International Publishing. doi: 10.1007/978-3-319-08613-2_30-1

Allen, J. J. B., Harmon-Jones, E., & Cavender, J. H. (2001). Manipulation of frontal EEG asymmetry through biofeedback alters self-reported emotional responses and facial EMG. *Psychophysiology*, *38*(4), 685–693. doi: 10.1111/1469-8986.3840685

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. doi: 10.1111/j.1469-8986.1997.tb02137.x

Amendola, K. L., & Wixted, J. T. (2015). No possibility of a selection bias, but direct evidence of a simultaneous superiority effect: a reply to Wells et al. *Journal of Experimental Criminology*, *11*(2), 291–294. doi: 10.1007/s11292-015-9227-x

Argaud, S., Delplanque, S., Houvenaghel, J. F., Auffret, M., Duprez, J., Vérin, M., … Sauleau, P. (2016). Does facial amimia impact the recognition of facial emotions? An EMG study in Parkinson's disease. *PLoS ONE*, *11*(7). doi:10.1371/journal.pone.0160329

Badham, S. P., Wade, K. A., Watts, H. J. E., Woods, N. G., & Maylor, E. A. (2013). Replicating distinctive facial features in lineups: Identification performance in young versus older adults. *Psychonomic Bulletin and Review*, *20*(2), 289–295. doi: 10.3758/s13423-012-0339-2

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin*, *133*(1), 1–24. doi: 10.1037/0033-2909.133.1.1

Barrett, S. E., Rugg, M. D., & Perrett, D. I. (1988). Event-related potentials and the matching of familiar and unfamiliar faces. *Neuropsychologia*, *26*(1), 105–117.

Beaudry, J. L., & Lindsay, R. C. L. (2006). Current identification procedure practices: A survey of Ontario police officers. *The Canadian Journal of Police and Security Services*, *4*(4), 178–183.

Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: an archival analysis. *Law and Human Behavior*, *25*, 475–491. doi: 10.1023/A:1012840831846

Bekerian, D. A., & Bowers, J. M. (1983). Eyewitness testimony: Were we misled? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(1), 139–145. doi: 10.1037/0278-7393.9.1.139

Berntson, G. G., Boysen, S. T., Bauer, H. R., & Torello, M. S. (1989). Conspecific screams and laughter: Cardiac and behavioral reactions of infant chimpanzees. *Developmental Psychobiology*, *22*(8), 771–787. doi: 10.1002/dev.420220803

Berntson, G. G., Quigley, K. S., Lozano, D., Cacioppo, J. T., & Tassinary, L. G. (2007). Cardiovascular psychophysiology. *Handbook of Psychophysiology*, 182–210. doi: 10.3109/07853890008995935

Blau, V. C., Maurer, U., Tottenham, N., & McCandliss, B. D. (2007). The face-specific N170 component is modulated by emotional facial expression. *Behavioral and Brain Functions*, *3*(1), 7. doi: 10.1186/1744-9081-3-7

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. In *Psychophysiology* (Vol. 46, pp. 1–11). doi: 10.1111/j.1469-8986.2008.00702.x

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276–298. doi: 10.1037/1528-3542.1.3.276

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. *Cognitive Neuroscience of Emotion*, *25*, 49–59.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602–607. doi: 10.1111/j.1469-8986.2008.00654.x

Bregman, N. J., & McAllister, H. A. (1982). Eyewitness Testimony: The Role of Commitment in Increasing Reliability. *Social Psychology Quarterly*. doi: 10.2307/3033652

Brewer, N., Caon, A., Todd, C., & Weber, N. (2006). Eyewitness identification accuracy and response latency. *Law and Human Behavior*, *30*, 31–50. doi: 10.1037/1076-898X.10.3.139

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied*, *12*(1), 11–30. doi: 10.1037/1076-898X.12.1.11

Brigham, J., Bennett, L., Meissner, C., & Mitchell, T. (2007). The influence of race on eyewitness memory. In D. Lindsay, D. Ross, D. Read, & M. Toglia (Eds.), *Handbook of Eyewitness Psychology Memory for People* (Vol. 2, pp. 257–281). Mahwah, NJ: Lawrence Erlbaum & Associates.

Brigham, J. C. (1999). What is forensic psychology, anyway? In *Law and Human Behavior* (Vol. 23, pp. 273–298). doi: 10.1023/A:1022304414537

Brigham, J. C., Maass, A., Snyder, L. D., & Spaulding, K. (1982). Accuracy of eyewitness identification in a field setting. *Journal of Personality and Social Psychology*, *42*(4), 673–681. doi: 10.1037/0022-3514.42.4.673

Brigham, J. C., & Ready, D. J. (1985). Own-race bias in lineup construction. *Law and Human Behavior*, *9*(4), 415–424. doi: 10.1007/BF01044480

Brigham, J. C., & Van Verst, M. (1986). Accuracy of Children's Eyewitness Identifications in a Field Setting. *Basic and Applied Social Psychology*, *7*(4), 295–306. doi: s15324834basp0704_4

Brown, E., Deffenbacher, K., & Sturgill, W. (1977). Memory for faces and the circumstances of encounter. *Journal of Applied Psychology*. doi: 10.1037/0021-9010.62.3.311

Brown, J. M. (2003). Eyewitness memory for arousing events: putting things into context. *Applied Cognitive Psychology*, *17*, 93–106. doi: 10.1002/acp.848

Brown, S. L., & Schwartz, G. E. (1980). Relationships between facial electromyography and subjective experience during affective imagery. *Biological Psychology*, *11*(1), 49–62. doi: 10.1016/0301-0511(80)90026-5

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305–327. doi: 10.1111/j.2044-8295.1986.tb02199.x

Buckhout, R. (1974). Eyewitness Testimony. *Scienticific American*, *231*, 23–31.

Buckhout, R., Alper, A., Chern, S., Silverberg, G., & Slomovits, M. (1974). Determinants of eyewitness performance on a lineup. *Bulletin of the Psychonomic Society*, *4*(3), 191–192. doi: 10.3758/BF03334241

Buhle, J. T., Silvers, J. A., Wage, T. D., Lopez, R., Onyemekwu, C., Kober, H., … Ochsner, K. N. (2014). Cognitive reappraisal of emotion: A meta-analysis of human neuroimaging studies. *Cerebral Cortex*, *24*(11), 2981–2990. doi: 10.1093/cercor/bht154

Busey, T. A., & Loftus, G. R. (2007). Cognitive science and the law. *Trends in Cognitive Sciences*. doi: 10.1016/j.tics.2006.12.004

Cacioppo, J. T., Martzke, J. S., Petty, R. E., & Tassinary, L. G. (1988). Specific forms of facial EMG response index emotions during an interview: from Darwin to the continuous flow hypothesis affect-laden information processing. *Journal of Personality and Social Psychology*, *54*(4), 592–604. doi: 10.1037/0022-

3514.54.4.592

Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, *50*(2), 260–268. doi: 10.1037/0022-3514.50.2.260

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (Eds.). (2007). *Handbook of psychophysiology*. Cambridge University Press. doi: 10.1017/CBO9780511546396

Caharel, S., Bernard, C., Thibaut, F., Haouzir, S., Di Maggio-Clozel, C., Allio, G., … Rebaï, M. (2007). The effects of familiarity and emotional expression on face processing examined by ERPs in patients with schizophrenia. *Schizophrenia Research*, *95*(1), 186–196. doi: 10.1016/j.schres.2007.06.015

Caharel, S., Poiroux, S., Bernard, C., Thibaut, F., Lalonde, R., & Rebai, M. (2002). Erps associated with familiarity and degree of familiarity during face recognition. *International Journal of Neuroscience*, *112*(12), 1499. doi: 10.1080/00207450290158368

Campbell, B. A., Wood, G., & McBride, T. (1997). Origins of orienting and defensive responses: An evolutionary perspective. *Attention and Orienting: Sensory and Motivational Processes*, 41–67.

Carlson, C. A., & Carlson, M. A. (2014). An evaluation of lineup presentation, weapon presence, and a distinctive feature using ROC analysis. *Journal of Applied Research in Memory and Cognition*, *3*(2), 45–53. doi: 10.1016/j.jarmac.2014.03.004

Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(4), 1228–1258. doi: 10.3758/s13415-014-0270-2

Carretié, L., Hinojosa, J. A., Martín-Loeches, M., Mercado, F., & Tapia, M. (2004). Automatic attention to emotional stimuli: Neural correlates. *Human Brain Mapping*, *22*(4), 290–299. doi: 10.1002/hbm.20037

Christianson, S. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin*, *112*(2), 284–309. doi: 10.1037/0033-2909.112.2.284

Christianson, S.-åke, & Loftus, E. F. (1987). Memory for traumatic events. *Applied Cognitive Psychology*, *1*, 225–239. doi: 10.1002/acp.2350010402

Clark, S. E., & Davey, S. L. (2005). The target-to-foils shift in simultaneous and sequential lineups. *Law and Human Behavior*, *29*, 151–172. doi: 10.1007/s10979-005-2418-7

Clifford, B. R., & Hollin, C. R. (1981). Effects of the type of incident and the number of perpetrators on eyewitness memory. *Journal of Applied Psychology*, *66*(3), 364–370. doi: 10.1037/0021-9010.66.3.364

Clifford, B. R., & Scott, J. (1978). Individual and situational factors in eyewitness testimony. *Journal of Applied Psychology*, *63*(3), 352–359. doi: 10.1037/0021-

9010.63.3.352

Colombo, J. (2001). The Development of Visual Attention in Infancy. *Annual Review of Psychology*, *52*(1), 337–367. doi: 10.1146/annurev.psych.52.1.337

Corey, D., Malpass, R. S., & McQuiston, D. E. (1999). Parallelism in eyewitness and mock witness identifications. *Applied Cognitive Psychology*, *13*(S1), S41–S58. doi: 10.1002/(SICI)1099-0720(199911)13:1+<S41::AID-ACP632>3.0.CO;2-A

Coxon, P., & Valentine, T. (1997). The effects of the age of eyewitnesses on the accuracy and suggestibility of their testimony. *Applied Cognitive Psychology*, *11*(5), 415–430. doi: 10.1002/(SICI)1099-0720(199710)11:5<415::AID-ACP469>3.0.CO;2-A

Cronin, C. (2006). *Forensic Psychology*. Kendall/Hunt Publishing Company.

Cumming, G. (2013). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. doi: 10.1177/0956797613504966

Cunningham, W. A., Van Bavel, J. J., Arbuckle, N. L., Packer, D. J., & Waggoner, A. S. (2012). Rapid social perception is flexible: approach and avoidance motivational states shape P100 responses to other-race faces. *Frontiers in Human Neuroscience*, *6*. doi: 10.3389/fnhum.2012.00140

Curley, L. J., MacLean, R., & Murray, J. (2017). The Relationship between the Big 5 Personality Traits and Eyewitness Recognition. *Journal of Articles in Support of the Null Hypothesis*, *13*(2).

Davidson, R. J. (1995). Cerebral asymmetry, emotion, and affective style. In Richard Davidson & Kenneth Hugdahl(Eds.) *Brain asymmetry* (pp. 361–387).London: The MIT Press

Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & Friesen, W. V. (1990). Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology: I. *Journal of Personality and Social Psychology*, *58*(2), 330–341. doi: 10.1037/0022-3514.58.2.330

Dawson, K., Hourihan, M. D., Wiles, C. M., & Chawla, J. C. (1994). Separation of voluntary and limbic activation of facial and respiratory muscles in ventral pontine infarction. *Journal of Neurology, Neurosurgery, and Psychiatry*, *57*(10), 1281–1282.

Deffenbacher, K. A. (1994). Effects of arousal on everyday memory. *Human Performance*, *7*(2), 141–161. doi: 10.1207/s15327043hup0702_3

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 and Human Behavior*. doi: 10.1007/s10979-004-0565-x

Delplanque, S., Silvert, L., Hot, P., & Sequeira, H. (2005). Event-related P3a and P3b in response to unpredictable emotional stimuli. *Biological Psychology*, *68*(2), 107–120. doi: 10.1016/j.biopsycho.2004.04.006

Denzel, S., & Possley, M. (2012). National Registry of EXONERATIONS - Terry Harrington

Case. Retrieved from https://www.law.umich.edu/special/exoneration/Pages/casedetail.aspx?caseid=3280

Devlin, P. (1976). *Report to the Secretary of State for the Home Department of the Departmental Committee on Evidence of Identification in Criminal Cases*. London: H.M. Stationery Off.

Dimberg, U. (1982). Facial reactions to facial expressions. *Psychophysiology*, *19*(6), 643–647. doi: 10.1111/j.1469-8986.1982.tb02516.x

Dimberg, U. (1990). Facial electromyography and emotional reactions. *Psychophysiology*, 27(5), 481-494. doi: 10.1111/j.1469-8986.1990.tb01962.x

Doob, A., & Kirshenbaum, H. (1973). Bias in police lineups — partial remembering. *Journal of Police Science and Administration*, *18*, 287–293.

Dror, I. E. (2012). Cognitive forensics and experimental research about bias in forensic casework. *Science and Justice*, *52*(2), 128-130. doi: 10.1016/j.scijus.2012.03.006

Dunning, D., & Perretta, S. (2002). Automaticity and eyewitness accuracy: A 10- to 12-second rule for distinguishing accurate from inaccurate positive identifications. *Journal of Applied Psychology*, *87*(5), 951–962. doi: 10.1037//0021-9010.87.5.951

Dunning, D., & Stern, L. B. (1994). Distinguishing accurate from inaccurate eyewitness identifications via inquiries about decision processes. *Journal of Personality and Social Psychology*, *67*(5), 818-835. doi: 10.1037/0022-3514.67.5.818

Dyer, A. G., Found, B., & Rogers, D. (2006). Visual attention and expertise for forensic signature analysis. *Journal of Forensic Sciences*, *51*(6), 1397–1404. doi: 10.1111/j.1556-4029.2006.00269.x

Dysart, J. E., Lawson, V. Z., & Rainey, A. (2012). Blind lineup administration as a prophylactic against the postidentification feedback effect. *Law and Human Behavior*, *36*(4), 312–319. doi: 10.1037/h0093921

Dysart, J., Lindsay, R., MacDonald, T., & Wicke, C. (2002). The intoxicated witness: Effects of alcohol on identification accuracy from showups. *Journal of Applied Psychology*, *87*(1), 170–175. doi: 10.1037/0021-9010.87.1.170

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport*, *11*(10), 2319–2324. doi: 10.1097/00001756-200007140-00050

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, *221*(4616), 1208–1210. doi: 10.1126/science.6612338

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception*, *8*(4), 431–439. doi: 10.1068/p080431

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1980). The deterioration of verbal descriptions of faces over different delay intervals. *Journal of Police Science and Administration*, *8*(1), 101–106.

Estrada-Reynolds, V. C., Reynolds, J. J., McCrea, S. M., & Freng, S. (2016). I Don't Like the Cut of your Jib: Perceived Facial Masculinity as a Cue to Criminality. *Psychiatry, Psychology and Law*, *24*(3), 392-409. doi: 10.1080/13218719.2016.1247639

Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy (''Lie Detection'') with Event-Related Brain Potentials. *Psychophysiology*, *28*(5), 531–547. doi: 10.1111/j.1469-8986.1991.tb01990.x

Feiling, A. (1927). Short Notes and Clinical Cases: A case of mimic facial paralysis. *Journal of Neurology and Psychopathology*, *8*(30), 141–145. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1068525/

Fisher, R. P., McCauley, M. R., & Geiselman, E. R. (1994). Improving eyewitness testimony with the cognitive interview. In *Adult Eyewitness Testimony: Current Trends and Developments*. Retrieved from http://books.google.com/books?hl=it&lr=lang_en%7Clang_it&id=GmWzZlp7imIC&pgis=1

Flowe, H. (2011). An exploration of visual behaviour in eyewitness identification tests. *Applied Cognitive Psychology*, *25*(2), 244–254. doi: 10.1002/acp.1670

Flowe, H., & Cottrell, G. (2011). An examination of simultaneous lineup identification decision processes using eye tracking. *Applied Cognitive Psychology*, *25*(3), 443–451. doi: 10.1002/acp.1711

Flowe, H. D. (2012). Do characteristics of faces that convey trustworthiness and dominance underlie perceptions of criminality? *PLoS ONE*, *7*(6), e37253-e37253. doi: 10.1371/journal.pone.0037253

Flowe, H. D., & Ebbesen, E. B. (2007). The effect of lineup member similarity on recognition accuracy in simultaneous and sequential lineups. *Law and Human Behavior*, *31*(1), 33–52. doi: 10.1007/s10979-006-9045-9

Flowe, H. D., & Humphries, J. E. (2011). An examination of criminal face bias in a random sample of police lineups. *Applied Cognitive Psychology*, *25*(2), 265–273. doi: 10.1002/acp.1673

Flowe, H. D., Mehta, A., & Ebbesen, E. B. (2011). The role of eyewitness identification evidence in felony case dispositions. *Psychology, Public Policy, and Law*, *17*(1), 140–159. doi: 10.1037/a0021311

Forensic neuroscience on trial (2001). *Nature Neuroscience*, *4*(1), 1–1. doi: 10.1038/82829

Forgas, J., Laham, S., & Vargas, P. (2005). Mood effects on eyewitness memory: Affective influences on susceptibility to misinformation. *Journal of Experimental Social Psychology*, *41*(6), 574–588. doi: 10.1016/j.jesp.2004.11.005

Foster, R. A., Libkuman, T. M., Schooler, J. W., & Loftus, E. F. (1994). Consequentiality and eyewitness person identification. *Applied Cognitive Psychology*, *8*(2), 107–121. doi: 10.1002/acp.2350080203

Friedman, B. H., & Thayer, J. F. (1998). Autonomic balance revisited: Panic anxiety and heart rate variability. *Journal of Psychosomatic Research*, 44(1), 133-151. doi: 10.1016/S0022-3999(97)00202-X

Gamer, M., Verschuere, B., Crombez, G., & Vossel, G. (2008). Combining physiological measures in the detection of concealed information. *Physiology & Behavior*, *95*(3), 333–340. doi: 10.1016/j.physbeh.2008.06.011

Gobbini, M. I., & Haxby, J. V. (2006). Neural response to the visual familiarity of faces. *Brain Research Bulletin*, *71*(1), 76–82. doi: 10.1016/j.brainresbull.2006.08.003

Gorenstein, G., & Ellsworth, P. (1980). Effect of choosing an incorrect photograph on a later identification by an eyewitness. *Journal of Applied Psychology*, *65*(5), 616–622. doi: 10.1037/0021-9010.65.5.616

Gothard, K. M. (2014). The amygdalo-motor pathways and the control of facial expressions. *Frontiers in Neuroscience*, *8*(43). doi: 10.3389/fnins.2014.00043

Greathouse, S. M., & Kovera, M. B. (2009). Instruction bias and lineup presentation moderate the effects of administrator knowledge on eyewitness identification. *Law and Human Behavior*, *33*(1), 70–82. doi: 10.1007/s10979-008-9136-x

Greene, E., & Evelo, A. J. (2014). Cops and robbers (and eyewitnesses): a comparison of lineup administration by robbery detectives in the USA and Canada. *Psychology, Crime and Law*, *21*(3), 297-313. doi: 10.1080/1068316X.2014.952236

Gronlund, S. D., Carlson, C. A., Dailey, S. B., & Goodsell, C. A. (2009). Robustness of the sequential lineup advantage. *Journal of Experimental Psychology. Applied*, *15*(2), 140–152. doi: 10.1037/a0015082

Gronlund, S. D., Carlson, C. A., Neuschatz, J. S., Goodsell, C. A., Wetmore, S. A., Wooten, A., & Graham, M. (2012). Showups versus lineups: An evaluation using ROC analysis. *Journal of Applied Research in Memory and Cognition*, *1*(4), 221–228. doi: 10.1016/j.jarmac.2012.09.003

Gronlund, S. D., Wixted, J. T., & Mickes, L. (2014). Evaluating Eyewitness Identification Procedures Using Receiver Operating Characteristic Analysis. *Current Directions in Psychological Science*, *23*(1), 3–10. doi: 10.1177/0963721413498891

Gross, S. R., & Shaffer, M. (2012). Exonerations in the United States, 1989–2012.

Hagsand, A., Roos-af-Hjelmsäter, E., Granhag, P. A., Fahlke, C., & Söderpalm-Gordh, A. (2013). Do sober eyewitnesses outperform alcohol intoxicated eyewitnesses in a lineup? *European Journal of Psychology Applied to Legal Context*, *5*(1), 23–47.

Hajcak, G., McDonald, N., & Simons, R. F. (2003). To err is autonomic: Error-related brain potentials, ANS activity, and post-error compensatory behavior. *Psychophysiology*, *40*(6), 895–903. doi: 10.1111/1469-8986.00107

Hajcak, G., McDonald, N., & Simons, R. F. (2004). Error-related psychophysiology and negative affect. *Brain and Cognition*, *56*(2 SPEC. ISS.), 189–197. doi: 10.1016/j.bandc.2003.11.001

Harris, M. J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin*, *97*(3), 363–386. doi: 10.1037/0033-2909.97.3.363

Hasan, H., Rehman, A. U., & Bhatti, M. I. (2016). New approach to an old issue related to ordered-choice endogenous models. *Quality & Quantity*, *50*(5), 2175-2184. doi: 10.1007/s11135-015-0257-2

Hastie, R., Landsman, R., & Loftus, E. F. (1978). Eyewitness testimony: The dangers of guessing. *Jurimetrics Journal*, 19, 1–8. doi: 10.2307/29761641

Havard, C., Memon, A., Laybourn, P., & Cunningham, C. (2012). Own-age bias in video lineups: A comparison between children and adults. *Psychology, Crime and Law*, *18*(10), 929–944. doi: 10.1080/1068316X.2011.598156

Haw, R., & Fisher, R. (2004). Effects of administrator-witness contact on eyewitness identification accuracy. *The Journal of Applied Psychology*, *89*(6), 1106–12. doi: 10.1037/0021-9010.89.6.1106

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition*, *33*(1), 98–106. doi: 10.3758/BF03195300

Henry, L. A., & Gudjonsson, G. H. (2007). Individual and developmental differences in eyewitness recall and suggestibility in children with intellectual disabilities. *Applied Cognitive Psychology*, *21*(3), 361–381. doi: 10.1002/acp

Herrmann, M. J., Ehlis, A.-C., Ellgring, H., & Fallgatter, A. J. (2005). Early stages (P100) of face perception in humans as measured with event-related potentials (ERPs). *Journal of Neural Transmission*, *112*(8), 1073–1081. doi: 10.1007/s00702-004-0250-8

Hilgendorf, E. L., & Irving, B. L. (1978). False Positive Identification. *Medicine, Science and the Law*, *18*(4), 255–262. doi: 10.1177/002580247801800407

Hope, L., Lewinski, W., Dixon, J., Blocksidge, D., & Gabbert, F. (2012). Witnesses in Action: The Effect of Physical Exertion on Recall and Recognition. *Psychological Science*, *23*(4), 386–390. doi: 10.1177/0956797611431463

Hosch, H. M., & Cooper, D. S. (1982). Victimization as a determinant of eyewitness accuracy. *Journal of Applied Psychology*, *67*(5), 649–652. doi: 10.1037/0021-9010.67.5.649

Hosch, H. M., Leippe, M. R., Marchioni, P. M., & Cooper, D. S. (1984). Victimization, self-monitoring, and eyewitness identification. *Journal of Applied Psychology*, *69*(2), 280–288. doi: 10.1037/0021-9010.69.2.280

Hot, P., Saito, Y., Mandai, O., Kobayashi, T., & Sequeira, H. (2006). An ERP investigation of emotional processing in European and Japanese individuals. *Brain Research*, *1122*(1), 171–178. doi: 10.1016/j.brainres.2006.09.020

Houston, K. A., Clifford, B. R., Phillips, L. H., & Memon, A. (2013). The emotional eyewitness: The effects of emotion on specific aspects of eyewitness recall and recognition performance. *Emotion*, *13*(1), 118–128. doi: 10.1037/a0029220

Hyman, R. (1989). The Psychology Of Deception. *Annual Review of Psychology*, *40*, 133–154. doi: 10.1146/annurev.psych.40.1.133

Iidaka, T., Harada, T., Kawaguchi, J., & Sadato, N. (2012). Neuroanatomical substrates involved in true and false memories for face. *NeuroImage*, *62*(1), 167–176. doi: 10.1016/j.neuroimage.2012.04.044

Iidaka, T., Harada, T., & Sadato, N. (2014). False memory for face in short-term memory and neural activity in human amygdala. *Brain Research*, *1591*(1), 74–85. doi: 10.1016/j.brainres.2014.10.003

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10), 1489–1506. doi: 10.1016/S0042-6989(99)00163-7

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259. doi: 10.1109/34.730558

Janik, S. W., Wellens, A. R., Goldberg, M. L., & Dell'Osso, L. F. (1978). Eyes as the center of focus in the visual examination of human faces. *Perceptual and Motor Skills*, *47*(3), 857–858. doi: 10.2466/pms.1978.47.3.857

Jemel, B., Pisani, M., Calabria, M., Crommelinck, M., & Bruyer, R. (2003). Is the N170 for faces cognitively penetrable? Evidence from repetition priming of Mooney faces of familiar and unfamiliar persons. *Cognitive Brain Research*, *17*(2), 431–446. doi: 10.1016/S0926-6410(03)00145-9

Johnsen, B. H., Thayer, J. F., & Hugdahl, K. (1995). Affective judgment of the Ekman faces: A dimensional approach. *Journal of Psychophysiology*, 9(3), 193-202.

Johnson, M., Hashtroudi, S., & Lindsay, D. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. doi: 10.1037/0033-2909.114.1.3

Jordan, D. (2005). Vagal control of the heart: central serotonergic (5-HT) mechanisms. *Experimental Physiology*, *90*(2), 175–181. doi: 10.1113/expphysiol.2004.029058

Kang, O. E., Huffer, K. E., & Wheatley, T. P. (2014). Pupil dilation dynamics track attention to high-level information. *PLoS ONE*, *9*(8), e102463. doi: 10.1371/journal.pone.0102463

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of*

Neuroscience, *17*(11), 4302–4311.

Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, *5*(2), 33–67. doi: 10.1111/j.1529-1006.2004.00016.x

Kassin, S. M., Tubb, V. a, Hosch, H. M., & Memon, a. (2001). On the "general acceptance" of eyewitness testimony research. A new survey of the experts. *The American Psychologist*, *56*, 405–416. doi: 10.1037/0003-066X.56.5.405

Kawamoto, T., Nittono, H., & Ura, M. (2013). Cognitive, affective, and motivational changes during ostracism: An ERP, EMG, and EEG study using a computerized Cyberball task. *Neuroscience Journal*, *2013*. doi:10.1155/2013/304674

Key, K. N., Cash, D. K., Neuschatz, J. S., Price, J., Wetmore, S. A., & Gronlund, S. D. (2015). Age differences (or lack thereof) in discriminability for lineups and showups. *Psychology, Crime and Law*, *21*(9), 871–889. doi: 10.1080/1068316X.2015.1054387

Klein Selle, N., Verschuere, B., Kindt, M., Meijer, E., & Ben-Shakhar, G. (2016). Orienting versus inhibition in the Concealed Information Test: Different cognitive processes drive different physiological measures. *Psychophysiology*, *53*(4), 579–590. doi: 10.1111/psyp.12583

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. doi: 10.1037/0022-3514.49.1.58

Künecke, J., Hildebrandt, A., Recio, G., Sommer, W., & Wilhelm, O. (2014). Facial EMG Responses to Emotional Expressions Are Related to Emotion Perception Ability. *PLOS ONE*, *9*(1), e84053. doi: 10.1371/journal.pone.0084053

Lacy, J. W., & Stark, C. E. L. (2013). The neuroscience of memory: Implications for the courtroom. *Nature Reviews Neuroscience*, *14*(9), 649–658. doi: 10.1038/nrn3563

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. In P. J. Lang, R. F. Simons, & M. T. Balaban (Eds.), *Attention and orienting: Sensory and motivational processes* (pp. 97-135). Mahwah, NJ: Lawrence Erlbaum Associates.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry*, *44*(12), 1248–1263. doi: 10.1016/S0006-3223(98)00275-3

Lapate, R. C., Rokers, B., Tromp, D. P. M., Orfali, N. S., Oler, J. A., Doran, S. T., … Davidson, R. J. (2016). Awareness of Emotional Stimuli Determines the Behavioral Consequences of Amygdala Activation and Amygdala-Prefrontal Connectivity. *Scientific Reports*, *6*(1), 25826. doi: 10.1038/srep25826

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, *40*(5), 776–785. doi: 10.1111/1469-8986.00078

Lefebvre, C. D., Marchand, Y., Smith, S. M., & Connolly, J. F. (2007). Determining eyewitness identification accuracy using event-related brain potentials (ERPs). *Psychophysiology*, *44*(6), 894–904. doi: 10.1111/j.1469-8986.2007.00566.x

Leippe, M. R., Wells, G. L., & Ostrom, T. M. (1978). Crime seriousness as a determinant of accuracy in eyewitness identification. *Journal of Applied Psychology*, *63*(3), 345–351. doi: 10.1037/0021-9010.63.3.345

Levett, L. M. (2013). Co-witness information influences whether a witness is likely to choose from a lineup. *Legal and Criminological Psychology*, *18*(1), 168–180. doi: 10.1111/j.2044-8333.2011.02033.x

Levi, A. (2016). Once again: selecting foils as similar to the suspect, or matching the description of the culprit? *Journal of Criminal Psychology*, *6*(3), 114–120. doi: 10.1108/JCP-03-2016-0011

Levi, A. M. (2007). Research note: Evidence for moving to an 84-person photo lineup. *Journal of Experimental Criminology*, *3*(4), 377–391. doi: 10.1007/s11292-007-9042-0

Levine, F. J., & Tapp, J. L. (1973). The psychology of criminal identification: The gap from Wade to Kirby. *University of Pennsylvania Law Review*, *121*(5), 1079–1131.

Lifshitz, K. (1966). The averaged evoked cortical response to complex visual stimuli. *Psychophysiology*, *3*(1), 55–68. doi: 10.1111/j.1469-8986.1966.tb02680.x

Lindsay, D. (1994). Memory source monitoring and eyewitness testimony. In D. Ross, D. Read, & M. Toglia (Eds.), *Adult eyewitness testimony: current trends and developments* (Vol. 5, pp. 27–55). New York: Cambridge University Press.

Lindsay, R. C., & Bellinger, K. (1999). Alternatives to the sequential lineup: The importance of controlling the pictures. *Journal of Applied Psychology*, *84*(3), 315–321. doi: 10.1037/0021-9010.84.3.315

Lindsay, R. C., Mansour, J. K., Beaudry, J. L., Leach, A.-M., & Bertrand, M. I. (2009a). Beyond sequential presentation: Misconceptions and misrepresentations of sequential lineups. *Legal and Criminological Psychology*, *14*(1), 31–34. doi: 10.1348/135532508X382104

Lindsay, R. C., Mansour, J. K., Beaudry, J. L., Leach, A.-M., & Bertrand, M. I. (2009b). Sequential lineup presentation: Patterns and policy. *Legal and Criminological Psychology*, *14*(1), 13–24. doi: 10.1348/135532508X382708

Lindsay, R. C., Martin, R., & Webber, L. (1994). Default values in eyewitness descriptions - A problem for the match-to-description lineup foil selection strategy. *Law and Human Behavior*, *18*(5), 527–541. doi: 10.1007/BF01499172

Lindsay, R. C., Pozzulo, J. D., Craig, W., Lee, K., & Corber, S. (1997). Simultaneous lineups, sequential lineups, and showups: Eyewitness identification decisions of adults and children. *Law and Human Behavior*, *21*(4), 391–404. doi: 10.1023/A:1024807202926

Lindsay, R. C., Semmler, C., Weber, N., Brewer, N., & Lindsay, M. R. (2008). How variations in distance affect eyewitness reports and identification accuracy. *Law and Human Behavior*, *32*(6), 526–535. doi: 10.1007/s10979-008-9128-x

Lindsay, R. C., Smith, S. M., & Pryke, S. (1999). Measures of Lineup Fairness: Do They Postdict Identification Accuracy? *Applied Cognitive Psychology*, *13*(1 SUPPL.), S93-S107. doi: 10.1002/(SICI)1099-0720(199911)13:1+3.3.CO;2-O

Lindsay, R. C., & Wells, G. L. (1980). What price justice? - Exploring the relationship of lineup fairness to identification accuracy. *Law and Human Behavior*, *4*(4), 303–313. doi: 10.1007/BF01040622

Lindsay, R. C., Lea, J. A., & Fulford, J. A. (1991). Sequential lineup presentation: Technique matters. *Journal of Applied Psychology*, *76*(5), 741-745. doi: 10.1037/0021-9010.76.5.741

Lindsay, R. C., Lea, J. A., Nosworthy, G. J., Fulford, J. A., Hector, J., LeVan, V., & Seabrook, C. (1991). Biased lineups: sequential presentation reduces the problem. *The Journal of Applied Psychology*, *76*(6), 796–802. https://doi.org/10.1037/0021-9010.76.6.796

Lindsay, R. C., & Wells, G. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology*, *70*(3), 556–564. doi: 10.1037/0021-9010.70.3.556

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: an MEG study. *Nature Neuroscience*, *5*(9), 910–916. doi:10.1038/nn909

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*(4), 560-572. doi: 10.1016/0010-0285(75)90023-7

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

Loftus, E. F., Loftus, G. R., & Messo, J. (1987). Some facts about “weapon focus.” *Law and Human Behavior*, *11*(1), 55–62. doi: 10.1007/BF01044839

Loftus, E. F., & Zanni, G. (1975). Eyewitness testimony: The influence of the wording of a question. *Bulletin of the Psychonomic Society*, *5*(1), 86–88. doi: 10.3758/BF03336715

Loftus, E. F., & Palmer, J. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*(5), 585-589. doi: 10.1016/S0022-5371(74)80011-3

Luque-Casado, A., Perales, J. C., Cárdenas, D., & Sanabria, D. (2016). Heart rate variability and cognitive processing: The autonomic response to task demands. *Biological Psychology*, *113*(Supplement C), 83–90. doi: 10.1016/j.biopsycho.2015.11.013

Luria, S. M., & Strauss, M. S. (1978). Comparison of eye movements over faces in photographic positives and negatives. *Perception*, *7*(3), 349–358. doi: 10.1068/p070349

Luus, C., & Wells, G. (1991). Eyewitness identification and the selection of distracters for lineups. *Law and Human Behavior*, *15*(1), 43–57. doi: 10.1007/BF01044829

MacLin, O., & MacLin, K. (2004). The effect of criminality on face attractiveness, typicality, memorability and recognition. *North American Journal of Psychology*, *6*(1), 145–154.

Malpass, R., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, *13*(4), 330–334. doi: 10.1037/h0028434

Malpass, R. S. (2006). A policy evaluation of simultaneous and sequential lineups. *Psychology, Public Policy, and Law*, 12(4), 394-418. doi: 10.1037/1076-8971.12.4.394

Malpass, R., & Tredoux, C. (2008). Lineup Size and Bias. In B. Cutler (Ed.), *Encyclopedia of Psychology and Law* (pp. 458–461). Thousand Oaks California: SAGE Publications, Inc. https://doi.org/10.4135/9781412959537

Mansour, J., & Flowe, H. (2010). Eyetracking and eyewitness memory. *Forensic Update*, *101*.

Mansour, J. K., Lindsay, R. C. L., Brewer, N., & Munhall, K. G. (2009). Characterizing visual behaviour in a lineup task. *Applied Cognitive Psychology*, *23*(7), 1012–1026. doi: 10.1002/acp.1570

Manzanero, A. L., Contreras, M. J., Recio, M., Alemany, A., & Martorell, A. (2012). Effects of presentation format and instructions on the ability of people with intellectual disability to identify faces. *Research in Developmental Disabilities*, *33*(2), 391–7. doi: 10.1016/j.ridd.2011.09.015

Marcon, J. L., Meissner, C. A., Jackiw, L. B., Arbuthnott, K. D., & Pfeifer, J. E. (2008). Examining the cross-race effect in lineup identification using caucasian and first nations samples. *Canadian Journal of Behavioural Science*, *40*(1), 52–57. doi: 10.1037/0008-400x.40.1.52

Massen v. State. 41 Wis. 2d 245 (1969).

Mavratzakis, A., Herbert, C., & Walla, P. (2016). Emotional facial expressions evoke faster orienting responses, but weaker emotional responses at neural and behavioural levels compared to scenes: A simultaneous EEG and facial EMG study. *NeuroImage*, *124*, 931–946. doi: 10.1016/j.neuroimage.2015.09.065

Mcllister, H. A., Michel, L. L. M., Tarcza, E. V, & Fitzmorris, J. M. (2008). Presentation procedures in lineups and mug books: A direct comparison. *Applied Cognitive Psychology*, *22*(2), 193–206. doi: 10.1002/acp.1370

McQuiston-Surrett, D., Malpass, R. S., & Tredoux, C. G. (2006). Sequential vs. Simultaneous Lineups: A Review of Methods, Data, and Theory. *Psychology, Public Policy, and Law*, *12*(2), 137-169. doi: 10.1037/1076-8971.12.2.137

Mecklenburg, S. H. (2006). Report to the legislature of the State of Illinois: The Illinois pilot program on double-blind, sequential lineup procedures. *Springfield, IL: Illinois State Police*.

Megreya, A. M., & Bindemann, M. (2013). Individual differences in personality and face identification. *Journal of Cognitive Psychology*, *25*(1), 30–37. doi: 10.1080/20445911.2012.739153

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: a dual-process signal detection theory analysis. *Memory & Cognition*, *33*(5), 783–792. doi: 10.3758/BF03193074

Meissner, C., Sporer, S., & Schooler, J. (2007). Person descriptions as eyewitness evidence. In D. Lindsay, D. Ross, D. Read, & M. Toglia (Eds.), *The handbook of eyewitness psychology, Vol II: Memory for people.* (pp. 3–34). Mahwah, NJ: Lawrence Erlbaum & Associates.

Memon, A., & Gabbert, F. (2003). Unravelling the effects of sequential presentation in culprit-present lineups. *Applied Cognitive Psychology*, *17*(6), 703–714. doi: 10.1002/acp.909

Memon, A., Hope, L., Bartlett, J., & Bull, R. (2002). Eyewitness recognition errors: the effects of mugshot viewing and choosing in young and old adults. *Memory & Cognition*, *30*(8), 1219–1227. doi: 10.3758/BF03213404

Mickes, L. (2015). Receiver operating characteristic analysis and confidence-accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory. *Journal of Applied Research in Memory and Cognition*, *4*(2), 93–102. doi: 10.1016/j.jarmac.2015.01.003

Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied*, *18*(4), 361–76. doi: 10.1037/a0030609

Mickes, L., Wais, P. E., & Wixted, J. T. (2009). Recollection is a continuous process: Implications for dual-process theories of recognition memory: Research Article. *Psychological Science*, *20*(4), 509–515. doi: 10.1111/j.1467-9280.2009.02324.x

Molinaro, P. F., Arndorfer, A., & Charman, S. D. (2013). Appearance-change instruction effects on eyewitness lineup identification accuracy are not moderated by amount of appearance change. *Law and Human Behavior*, *37*(6), 432–440. doi: 10.1037/lhb0000049

Münsterberg, H. (1908). *On the Witness Stand: Essays on Psychology and Crime*. New York: Doubleday.

Münsterberg, M. (1922). *Hugo Münsterberg, his life and work*. New York: D. Appleton & Co.

Murray, D. M., & Wells, G. L. (1982). Does Knowledge that a Crime Was Staged Affect Eyewitness Performance? *Journal of Applied Social Psychology*, *12*(1), 42–53. doi:

10.1111/j.1559-1816.1982.tb00847.x

Mushiake, H., Saito, N., Sakamoto, K., Itoyama, Y., & Tanji, J. (2006). Activity in the Lateral Prefrontal Cortex Reflects Multiple Steps of Future Events in Action Plans. *Neuron*, *50*(4), 631–641. doi: 10.1016/j.neuron.2006.03.045

Naber, M., & Nakayama, K. (2013). Pupil responses to high-level image content. *Journal of Vision*, *13*(6), 1-8. doi:10.1167/13.6.7

Narby, D., Cutler, B., & Penrod, S. (1996). The effects of witness, target, and situational factors on eyewitness identifications. In S. Sporer, R. Malpass, & G. Koehnken (Eds.), *Psychological issues in eyewitness identification.* (pp. 23–52). Mahwah, NJ: Lawrence Erlbaum.

Nosworthy, G., & Lindsay, R. (1990). Does nominal lineup size matter? *Journal of Applied Psychology*, *75*(3), 358–361. doi: 10.1037/0021-9010.75.3.358

O'Connell, M., & Synnott, J. (2009). A position of influence: Variation in offender identification rates by location in a lineup. *Journal of Investigative Psychology and Offender Profiling*, *6*(2), 139–149. doi: 10.1002/jip.102

Olivares, E. I., & Iglesias, J. (2008). Brain potentials and integration of external and internal features into face representations. *International Journal of Psychophysiology*, *68*(1), 59–69. doi: 10.1016/j.ijpsycho.2008.01.003

Olofsson, J. K., & Polich, J. (2007). Affective visual event-related potentials: Arousal, repetition, and time-on-task. *Biological Psychology*, *75*(1), 101–108. doi: 10.1016/j.biopsycho.2006.12.006

Palmer, M. A., Sauer, J. D., & Holt, G. A. (2017). Undermining position effects in choices from arrays, with implications for police lineups. *Journal of Experimental Psychology: Applied*, *23*(1), 71–84. doi: 10.1037/xap0000109

Palomba, D., Angrilli, A., & Mini, A. (1997). Visual evoked potentials, heart rate responses and memory to emotional pictorial stimuli. *International Journal of Psychophysiology*, *27*(1), 55–67. doi: 10.1016/S0167-8760(97)00751-4

Pan, Q., Zhou, G., Wang, R., Cai, G., Yan, J., Fang, L., & Ning, G. (2016). Do the deceleration/acceleration capacities of heart rate reflect cardiac sympathetic or vagal activity? A model study. *Medical & Biological Engineering & Computing*, *54*(12), 1921–1933. doi: 10.1007/s11517-016-1486-9

Parker, J. F., & Ryan, V. (1993). An attempt to reduce guessing behavior in children's and adults' eyewitness identifications. *Law and Human Behavior, 27*(1), 55-67. doi: 10.1007/BF01044534

Pascual-Marqui, R. D., Esslen, M., Kochi, K., & Lehmann, D. (2002). Functional imaging with low-resolution brain electromagnetic tomography (LORETA): A review. *Methods and Findings in Experimental and Clinical Pharmacology*, *24*(Suppl C), 91–5. https://doi.org/841 [pii]

Pascual-Marqui, R. D., Lehmann, D., Koenig, T., Kochi, K., Merlo, M. C. G., Hell, D., & Koukkou, M. (1999). Low resolution brain electromagnetic tomography (LORETA) functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia. *Psychiatry Research: Neuroimaging*, *90*(3), 169–179. doi: 10.1016/S0925-4927(99)00013-X

Penrod, S., & Bornstein, B. (2007). Generalizing Eyewitness. Reliability Research. In R. Lindsay, D. Ross, D. Read, & M. Toglia (Eds.), *Handbook of Eyewitness Psychology Memory for People* (2nd ed., pp. 529–556). Mahwah, NJ: Erlbaum.

Penrod, S., Loftus, E., & Winkler, J. (1982). The reliability of eyewitness testimony: A psychological perspective. In N. Kerr & R. Bray (Eds.), *The psychology of the courtroom*. New York: Academic Press.

People v. Chambers. 112 Ill. App. 2d 347 (1969).

People v. Stanton. 274 Cal. App. 2d 13 (1969).

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: from a "low road" to "many roads" of evaluating biological significance. *Nature Reviews Neuroscience*, *11*(11), 773–783. doi: 10.1038/nrn2920

Phillips, M. L., Williams, L. M., Heining, M., Herba, C. M., Russell, T., Andrew, C., … Gray, J. A. (2004). Differential neural responses to overt and covert presentations of facial expressions of fear and disgust. *NeuroImage*, *21*(4), 1484–1496. doi: 10.1016/j.neuroimage.2003.12.013

Picard, N., & Strick, P. L. (2001). Imaging the premotor areas. *Current Opinion in Neurobiology*, *11*(6), 663-672. doi: 10.1016/S0959-4388(01)00266-5

Poole, A., & Ball, L. J. (2005). Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects. *Encyclopedia of Human-Computer Interaction*, 211–219. https://doi.org/10.4018/978-1-59140-562-7

Porges, S. W. (1992). Autonomic regulation and attention. In *Attention and information processing in infants and adults: perspectives from human and animal research* (pp. 201–223).

Pozzulo, J. D., Dempsey, J. L., & Wells, K. (2010). Does Lineup Size Matter with Child Witnesses. *Journal of Police and Criminal Psychology*, *25*(1), 22–26. doi: 10.1007/s11896-009-9055-x

Pozzulo, J. D., & Marciniak, S. (2006). Comparing identification procedures when the perpetrator has changed appearance. *Psychology, Crime and Law*, *12*(4), 429–438. doi: 10.1080/10683160500050690

Pozzulo, J. D., Reed, J., Pettalia, J., & Dempsey, J. (2016). Simultaneous, Sequential, Elimination, and Wildcard: A Comparison of Lineup Procedures. *Journal of Police and Criminal Psychology*, *31*(1), 71–80. doi: 10.1007/s11896-015-9168-3

Pozzulo, J. D., & Warren, K. L. (2003). Descriptions and identifications of strangers by youth and adult eyewitnesses. *The Journal of Applied Psychology*, *88*(2), 315–323. https://doi.org/10.1037/0021-9010.88.2.315

Pozzulo, J., Dempsey, J., Crescini, C., & Lemieux, J. (2009). Examining the relation between eyewitness recall and recognition for children and adults. *Psychology, Crime & Law*, *15*(5), 409–424. doi: 10.1080/10683160802279625

Quintana, D. S., Guastella, A. J., Outhred, T., Hickie, I. B., & Kemp, A. H. (2012). Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the autonomic nervous system and social cognition. *International Journal of Psychophysiology*, *86*(2), 168–172. doi: 10.1016/j.ijpsycho.2012.08.012

Reed, T. E. (1984). Eyewitness identifications of an armed robber within a biased police lineup. *Journal of Police Science and Administration*, *12*(3), 310–314.

Reinders, A., Gläscher, J., De Jong, J. R., Willemsen, A. T. M., Den Boer, J. A., & Büchel, C. (2006). Detecting fearful and neutral faces: BOLD latency differences in amygdala–hippocampal junction. *Neuroimage*, *33*(2), 805–814. doi: 10.1016/j.neuroimage.2006.06.052

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, *7*(1), 1–75. doi: 10.1016/1041-6080(95)90031-4

Ribeiro, R. L., Teixeira-Silva, F., Pompéia, S., & Bueno, O. F. A. (2007). IAPS includes photographs that elicit low-arousal physiological responses in healthy volunteers. *Physiology and Behavior*, *91*(5), 671–675. doi: 10.1016/j.physbeh.2007.03.031

Richards, J. E., & Casey, B. J. (1992). Development of sustained visual attention in the human infant. In B. A. Campbell, H. Hayne, & R. Richardson (Eds.), *Attention and information processing in infants and adults: Perspectives from human and animal research* (pp. 30-60). Hillsdale, NJ: Lawrence Erlbaum Associates.

Righi, S., Marzi, T., Toscani, M., Baldassi, S., Ottonello, S., & Viggiano, M. P. (2012). Fearful expressions enhance recognition memory: Electrophysiological evidence. *Acta Psychologica*, *139*(1), 7–18. doi: 10.1016/j.actpsy.2011.09.015

Rogers, R. F., Rybak, I. A., & Schwaber, J. S. (2000). Computational modeling of the baroreflex arc: Nucleus tractus solitarius. *Brain Research Bulletin*, *51*(2), 139–150. doi: 10.1016/S0361-9230(99)00242-7

Rose, C., & Beck, V. (2016). Eyewitness accounts: false facts, false memories, and false identification. *Journal of Crime and Justice*, *39*(2), 243–263. doi: 10.1080/0735648X.2014.940999

Rose, R. A., Bull, R., & Vrij, A. (2005). Non-biased lineup instructions do matter - A problem for older witnesses. *Psychology, Crime and Law*, *11*(2), 147–159. doi: 10.1080/10683160512331316307

Rosenfeld, J. P. (2002). Event-related potentials in detection of deception,malingering, and false memories. In M. Kleiner (Ed.), *Handbook of Polygraph Testing* New York: Academic Press.

Rossion, B., Hanseeuw, B., & Dricot, L. (2012). Defining face perception areas in the human brain: a large-scale factorial fMRI face localizer analysis. *Brain and Cognition*, *79*(2), 138–157. doi: 10.1016/j.bandc.2012.01.001

Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage*, *39*(4), 1959–1979. doi:10.1016/j.neuroimage.2007.10.011

Safer, M., Christianson, S., Autry, M., & Österlund, K. (1998). Tunnel memory for traumatic events. *Applied Cognitive Psychology*, *12*(2), 99–117. doi: 10.1002/(SICI)1099-0720(199804)12:2<99::AID-ACP509>3.3.CO;2-Z

Saladin, M., Saper, Z., & Breen, L. (1988). Perceived attractiveness and attributions of criminality: What is beautiful is not criminal. *Canadian Journal of Criminology*, *30*(3), 251–259.

Santos, I. M., Soares, S. C., Oliveira, B., Alho, L., Bem-Haja, P., & Silva, C. F. (2014). Contributions of experimental psychology to forensic science. In A. M. Columbus (Ed.), *Advances in Psychology Research* (Vol. 100, pp. 25-58). New York: Nova Science Publishers. ISBN: 978-1-63321-482-8

Schwartz, G. E., Fair, P. L., Salt, P., Mandel, M. R., & Klerman, G. L. (1976). Facial muscle patterning to affective imagery in depressed and nondepressed subjects. *Science*, *192*(4238), 489–491. doi: 10.1126/science.1257786

Sestito, M., Umiltà, M. A., De Paola, G., Fortunati, R., Raballo, A., Leuci, E., … Gallese, V. (2013). Facial reactions in response to dynamic emotional stimuli in different modalities in patients suffering from schizophrenia: a behavioral and EMG study. *Frontiers in Human Neuroscience*, *7*, 368. doi: 10.3389/fnhum.2013.00368

Shepherd, J. W., Davies, G. M., & Ellis, H. D. (1978). How best shall a face be described. *Practical Aspects of Memory*, 278–285.

Shibata, T., Nishijo, H., Tamura, R., Miyamoto, K., Eifuku, S., Endo, S., & Ono, T. (2002). Generators of visual evoked potentials for faces and eyes in the human brain as determined by dipole localization. *Brain Topography*, *15*(1), 51–63. doi: 10.1023/A:1019944607316

Simola, J., Torniainen, J., Moisala, M., Kivikangas, M., & Krause, C. M. (2013). Eye movement related brain responses to emotional scenes during free viewing. *Frontiers in Systems Neuroscience*, *7(41)*. doi: 10.3389/fnsys.2013.00041

Smith, A. M., & Cutler, B. L. (2013). Introduction: Identification procedures and conviction of the innocent. In: Cutler BL (ed.), Reform of Eyewitness Identification Procedures. Washington, DC: American Psychological Association

Smith, S. M., Lindsay, R. C. L., & Pryke, S. (2000). Postdictors of eyewitness errors: Can false identifications be diagnosed? *Journal of Applied Psychology*, *85*(4), 542–550. doi: 10.1037/0021-9010.85.4.542

Smith, S. M., Stinson, V., & Prosser, M. A. (2004). Do they all look alike? An exploration of decision-making strategies in cross-race facial identifications. *Canadian Journal of Behavioural Science*, *36*(2), 146–154. doi: 10.1037/h0087225

Smith, D. V., Clithero, J. A., Boltuck, S. E., & Huettel, S. A. (2013). Functional connectivity with ventromedial prefrontal cortex reflects subjective value for social rewards. *Social Cognitive and Affective Neuroscience*, *9*(12), 2017–2025. doi: 10.1093/scan/nsu005

Soares, S. C., Maior, R. S., Isbell, L. A., Tomaz, C., & Nishijo, H. (2017). Fast Detector/First Responder: Interactions between the Superior Colliculus-Pulvinar Pathway and Stimuli Relevant to Primates. *Frontiers in Human Neuroscience*, 11(67) doi: 10.3389/fnins.2017.00067

Sokolov, E. N. (1963). Higher nervous functions: The orienting reflex. *Annual Review of Physiology*, *25*(1), 545–580. doi: 10.1146/annurev.ph.25.030163.002553

Stacey, P. C., Walker, S., & Underwood, J. D. M. (2005). Face processing and familiarity: Evidence from eye-movement data. *British Journal of Psychology*, *96*(4), 407–422. doi: 10.1348/000712605X47422

State v. Burch. (1969). 284 Minn. 300.

State v. Parker.(1969). 282 Minn. 343.

Steblay, N. (2006). Reforming eyewitness identification: Lineup identification instructions; weighing the advantages and disadvantages of show-ups versus lineups. *Cardozo Public Law, Policy & Ethics Journal*, *4*(2), 341–354.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, *25*(5), 459–473.

Steblay, N. K., Dietrich, H. L., Ryan, S. L., Raczynski, J. L., & James, K. A. (2011). Sequential lineup laps and eyewitness accuracy. *Law and Human Behavior*, *35*, 262–274. doi: 10.1007/s10979-010-9236-2

Steblay, N. K., Dysart, J. E., & Wells, G. L. (2011). Seventy-two tests of the sequential lineup superiority effect: A meta-analysis and policy discussion. *Psychology, Public Policy, and Law*, *17*(1), 99–139. doi: 10.1037/a0021650

Stovall v. Denno. (1967). 388 U.S. 263.

Synnott, J., Dietzel, D., & Ioannou, M. (2015). A review of the polygraph: history, methodology and current status. *Crime Psychology Review*, *1*(1), 59–83. doi: 10.1080/23744006.2015.1060080

Tamietto, M., Pullens, P., De Gelder, B., Weiskrantz, L., & Goebel, R. (2012). Subcortical connections to human amygdala and changes following destruction of the visual cortex. *Current Biology*, *22*(15), 1449–1455. doi: 10.1016/j.cub.2012.06.006

Tan, J.W., Walter, S., Scheck, A., Hrabal, D., Hoffmann, H., Kessler, H., & Traue, H. C. (2012). Repeatability of facial electromyography (EMG) activity over corrugator supercilii and zygomaticus major on differentiating various emotions. *Journal of Ambient Intelligence and Humanized Computing*, *3*(1), 3–10. doi: 10.1007/s12652-011-0084-9

Tan, J. W., Andrade, J. W., Li, H., Walter, S., Hrabal, D., Rukavina, S., … Traue, H. C. (2016). Recognition of intensive valence and arousal affective states via facial electromyographic activity in young and senior adults. *PLoS ONE*, *11*(1). doi: 10.1371/journal.pone.0146691

Ternes, M., & Yuille, J. C. (2008). Eyewitness Memory and Eyewitness Identification Performance in Adults with Intellectual Disabilities. *Journal of Applied Research in Intellectual Disabilities*, *21*(6), 519–531. doi: 10.1111/j.1468-3148.2008.00425.x

The Innocence Project. (2017). Innocent Project case profile. Retrieved from http:/www.innocenceproject.org/know/

Tredoux, C., Parker, J. F., & Nunez, D. (2007). Predicting Eyewitness Identification Accuracy with Mock Witness Measures of Lineup Fairness: Quality of Encoding Interacts with Lineup Format. *South African Journal of Psychology*, *37*, 207–222. doi: 10.1177/008124630703700201

Trepel, M., Weller, M., Dichgans, J., & Petersen, D. (1996). Voluntary facial palsy with a pontine lesion. *Journal of Neurology, Neurosurgery, and Psychiatry*, *61*(5), 531–533.

Tromp, D. P. M., Grupe, D. W., Oathes, D. J., McFarlin, D. R., Hernandez, P. J., Kral, T. R. A., … Nitschke, J. B. (2012). Reduced Structural Connectivity of a Major Frontolimbic Pathway in Generalized Anxiety Disorder. *Archives of General Psychiatry*, *69*(9), 925. doi: 10.1001/archgenpsychiatry.2011.2178

Underwood, M. D., Arango, V., Bakalian, M. J., Ruggiero, D. A., & Mann, J. J. (1999). Dorsal raphe nucleus serotonergic neurons innervate the rostral ventrolateral medulla in rat. *Brain Research*, *824*(1), 45–55. doi: 10.1016/S0006-8993(99)01181-6

Ursino, M., & Magosso, E. (2003). Role of short-term cardiovascular regulation in heart period variability: a modeling study. *American Journal of Physiology - Heart and Circulatory Physiology*, *284*(4), H1479–H1493. doi: 10.1152/ajpheart.00850.2002

Valentine, T., & Mesout, J. (2009). Eyewitness identification under stress in the London dungeon. *Applied Cognitive Psychology*, *23*, 151–161. doi: 10.1002/acp.1463

Van Kesteren, M. T. R., & Brown, T. I. (2014). The Medial Prefrontal Cortex and the Deceptiveness of Memory. *The Journal of Neuroscience*, *34*(41), 13569 -13570. doi: 10.1523/JNEUROSCI.3091-14.2014

Vernet, M., Quentin, R., Chanes, L., Mitsumasu, A., & Valero-Cabre, A. (2014).

Corrigendum: Frontal eye field, where art thou? Anatomy, function, and non-invasive manipulation of frontal regions involved in eye movements and associated cognitive operations. *Frontiers in Integrative Neuroscience*, *8*, 88. doi: 10.3389/fnint.2014.00088

Verschuere, B., Crombez, G., Koster, E. H. W., Van Bockstaele, B., & De Clercq, A. (2007). Startling secrets: Startle eye blink modulation by concealed crime information. *Biological Psychology*, *76*(1), 52–60. doi: 10.1016/j.biopsycho.2007.06.001

Verschuere, B., Crombez, G., Smolders, L., & De Clercq, A. (2009). Differentiating orienting and defensive responses to concealed information: The role of verbalization. *Applied Psychophysiology and Biofeedback*, *34*(3), 237–244. doi: 10.1007/s10484-009-9093-2

Vidal, J. R., Ossandón, T., Jerbi, K., Dalal, S. S., Minotti, L., Ryvlin, P., … Lachaux, J.-P. (2010). Category-Specific Visual Responses: An Intracranial Study Comparing Gamma, Beta, Alpha, and ERP Response Selectivity. *Frontiers in Human Neuroscience*, *4*, 195. doi: 10.3389/fnhum.2010.00195

Von Der Heide, R. J., Skipper, L. M., & Olson, I. R. (2013). Anterior temporal face patches: a meta-analysis and empirical study. *Frontiers in Human Neuroscience*, *7*. doi: 10.3389/fnhum.2013.00017

Vrij, A., Mann, S., & Fisher, R. P. (2006). Information-gathering vs accusatory interview style: Individual differences in respondents' experiences. *Personality and Individual Differences*, *41*(4), 589–599. doi: 10.1016/j.paid.2006.02.014

Wagenaar, W. A., & Van Der Schrier, J. H. (1996). Face recognition as a function of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime and Law*, *2*(4), 321–332. doi: 10.1080/10683169608409787

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks*, *19*(9), 1395–1407. doi: 10.1016/j.neunet.2006.10.001

Wardak, C., Ibos, G., Duhamel, J.-R., & Olivier, E. (2006). Contribution of the monkey frontal eye field to covert visual attention. *Journal of Neuroscience*, *26*(16), 4228–4235. doi: 10.1523/JNEUROSCI.3336-05.2006

Warnick, D. H., & Sanders, G. S. (1980a). The Effects of Group Discussion on Eyewitness Accuracy. *Journal of Applied Social Psychology*, *10*(3), 249–259. doi: 10.1111/j.1559-1816.1980.tb00707.x

Warnick, D. H., & Sanders, G. S. (1980b). Why Do Eyewitnesses Make So Many Mistakes? *Journal of Applied Social Psychology*, *10*(4), 362–366. doi: 10.1111/j.1559-1816.1980.tb00716.x

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), 139–147. doi: 10.1037/1076-898X.10.3.139

Weinberg, H. I., Wadsworth, J., & Baron, R. S. (1983). Demand and the impact of leading questions on eyewitness testimony. *Memory & Cognition*, *11*(1), 101–104. doi: 10.3758/BF03197667

Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, *36*(12), 1546–1557. doi: 10.1037/0022-3514.36.12.1546

Wells, G. L. (1984). The Psychology of Lineup Identifications. *Journal of Applied Social Psychology*, *14*(2), 89–103. doi: 10.1111/j.1559-1816.1984.tb02223.x

Wells, G. L. (2014). Eyewitness Identification: Probative Value, Criterion Shifts, and Policy Regarding the Sequential Lineup. *Current Directions in Psychological Science*, *23*(1), 11–16. doi: 10.1177/0963721413504781

Wells, G. L., & Bradfield, A. L. (1999). Measuring the Goodness of Lineups: Parameter Estimation, Question Effects, and Limits to the Mock Witness Paradigm. *Applied Cognitive Psychology*, *13*(1 SUPPL.). doi: 10.1002/(SICI)1099-0720(199911)13:1+3.3.CO;2-D

Wells, G. L., Dysart, J. E., & Steblay, N. K. (2015). The flaw in Amendola and Wixted's conclusion on simultaneous versus sequential lineups. *Journal of Experimental Criminology*, *11*(2), 285–289. doi: 10.1007/s11292-014-9225-4

Wells, G. L., & Elizabeth Luus, C. A. (1990). Police lineups as experiments: Social methodology as a framework for properly conducted lineups. *Personality and Social Psychology Bulletin*, *16*(1), 106–117. doi: 10.1177/0146167290161008

Wells, G. L., Malpass, R. O. Y. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., Fulero, S. M., … Todd, C. (2009). Divergent viewpoints about lineups: What policy for the identification of criminals? *Law and Human Behavior*, *14*(5), 257–265. doi: 10.1016/j.erap.2013.07.008

Wells, G. L., Memon, A., & Penrod, S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest*, *7*, 45–75. doi: 10.1111/j.1529-1006.2006.00027.x

Wells, G. L., & Olson, E. A. (2003). Eyewitness Testimony. *Annual Review of Psychology*, *54*(1), 277–295. doi: 10.1146/annurev.psych.54.101601.145028

Wells, G. L., Leippe, M., & Ostrom, T. (1979). Guidelines for empirically assessing the fairness of a lineup. *Law and Human Behavior*, *3*(4), 285–293. doi: 10.1007/BF01039807

Wells, G. L., Lindsay, R., & Ferguson, T. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *The Journal of Applied Psychology*, *64*(4), 440–448. doi: 10.1037/0021-9010.64.4.440

Wells, G. L., Seelau, E., Rydell, S., & Luus, E. (1994). Recommendations for properly conducted lineup identification tasks. In D. Ross, D. Read, & M. Toglia (Eds.), *Adult Eyewitness Testimony: Current Trends and Developments* (pp. 223–244). New York:

Cambridge University Press.

Werner, N. S., Kühnel, S., & Markowitsch, H. J. (2013). The Neuroscience of Face Processing and Identification in Eyewitnesses and Offenders. *Frontiers in Behavioral Neuroscience*, *7*, 189. doi: 10.3389/fnbeh.2013.00189

Westlye, L. T., Bjørnebekk, A., Grydeland, H., Fjell, A. M., & Walhovd, K. B. (2011). Linking an anxiety-related personality trait to brain white matter microstructure: Diffusion tensor imaging and harm avoidance. *Archives of General Psychiatry*, *68*(4), 369–377. doi: 10.1001/archgenpsychiatry.2011.24

Whipple, G. (1909). The observer as reporter: A survey of the “psychology of testimony”. *Psychological Bulletin*, *6*(5), 153–170. doi: 10.1037/h0071084

Winecoff, A., Clithero, J. A., Carter, R. M., Bergman, S. R., Wang, L., & Huettel, S. A. (2013). Ventromedial Prefrontal Cortex Encodes Emotional Value. *The Journal of Neuroscience*, *33*(27), 11032–11039. doi: 10.1523/JNEUROSCI.4317-12.2013

Wise, R., Cushman, B., & Safer, M. (2012). Scientific Procedures for Eyewitness Evidence for U.S. Law Officers. *Journal of Forensic Psychology Practice*, *12*(2), 173–187. doi: 10.1080/15228932.2012.650145

Wise, R., Dauphinais, K., & Safer, M. (2007). A Tripartite Solution to Eyewitness Error. *The Journal of Criminal Law and Criminology*, *97*(1), 807–872. doi: 10.2307/40042843

Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychological Review*, *121*(2), 262–276. doi: 10.1037/a0035940

Wixted, J. T., & Mickes, L. (2015). Evaluating eyewitness identification procedures: ROC analysis and its misconceptions. *Journal of Applied Research in Memory and Cognition*, *4*(4), 318–323. doi: 10.1016/j.jarmac.2015.08.009

Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, *18*(1), 10–65. doi: 10.1177/1529100616686966

Wogalter, M. S., Malpass, R. S., & Mcquiston, D. E. (2004). A national survey of us police on preparation and conduct of identification lineups. *Psychology, Crime & Law*, *10*(1), 69-82. doi: 10.1080/10683160410001641873

Wylie, L. E., Bergt, S., Haby, J., Brank, E. M., & Bornstein, B. H. (2015). Age and lineup type differences in the own-race bias. *Psychology, Crime and Law*, *21*(5), 490–506. doi: 10.1080/1068316X.2014.989173

Yarmey, A. (1993). Stereotypes and recognition memory for faces and voices of good guys and bad guys. *Applied Cognitive Psychology*, *7*(5), 419–431. doi: 10.1002/acp.2350070505

Yarmey, A. D., & Morris, S. (1998). The effects of discussion on eyewitness memory.

Journal of Applied Social Psychology, *28*(17), 1637–1648. doi: 10.1111/j.1559-1816.1998.tb01693.x

Yarmey, A. D., Yarmey, M. J., & Yarmey, A. L. (1996). Accuracy of eyewitness identifications in showups and lineups. *Law and Human Behavior*. doi: 10.1007/BF01498981

Yonelinas, A. P., Aly, M., Wang, W. C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194. doi: 10.1002/hipo.20864

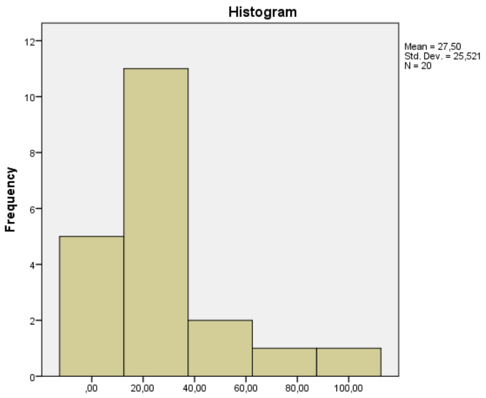
Young, A. W., Hay, D. C., McWeeny, K. H., Flude, B. M., & Ellis, A. W. (1985). Matching familiar and unfamiliar faces on internal and external features. *Perception*, *14*(6), 737–746. doi: 10.1068/p140737

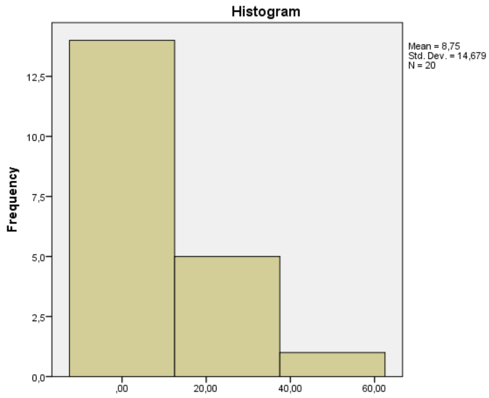
Zaitsu, W. (2016). External validity of Concealed Information Test experiment: Comparison of respiration, skin conductance, and heart rate between experimental and field card tests. *Psychophysiology*, *53*(7), 1100–1107. doi: 10.1111/psyp.12650

APPENDIX

Appendix - CRV_ Analysis: Example for Misses

SOURCE – ORIGINAL DATA

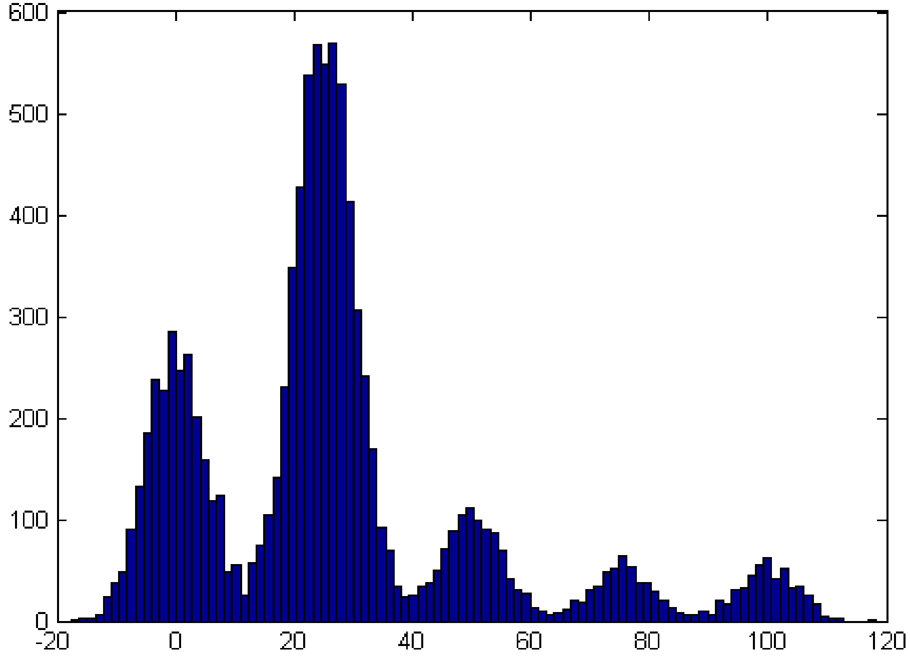


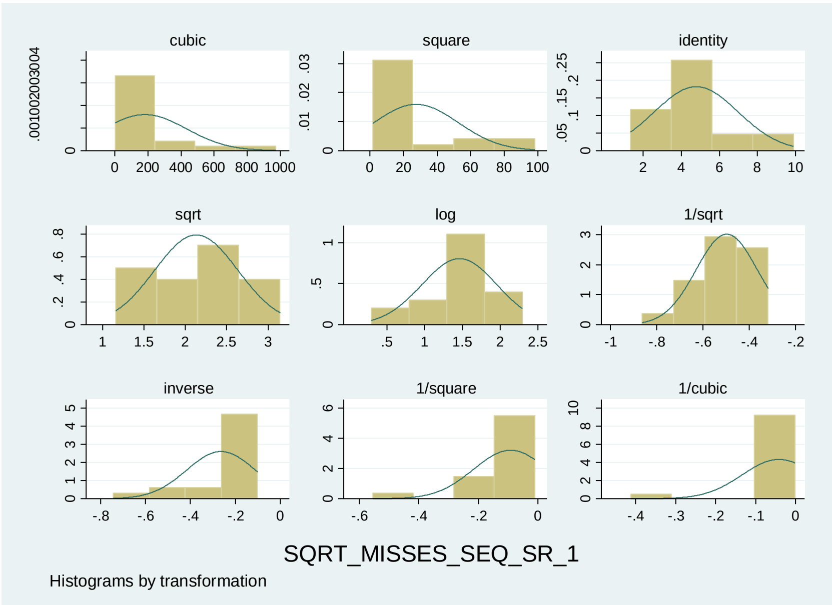


1 - Matlab Script - applied to original data

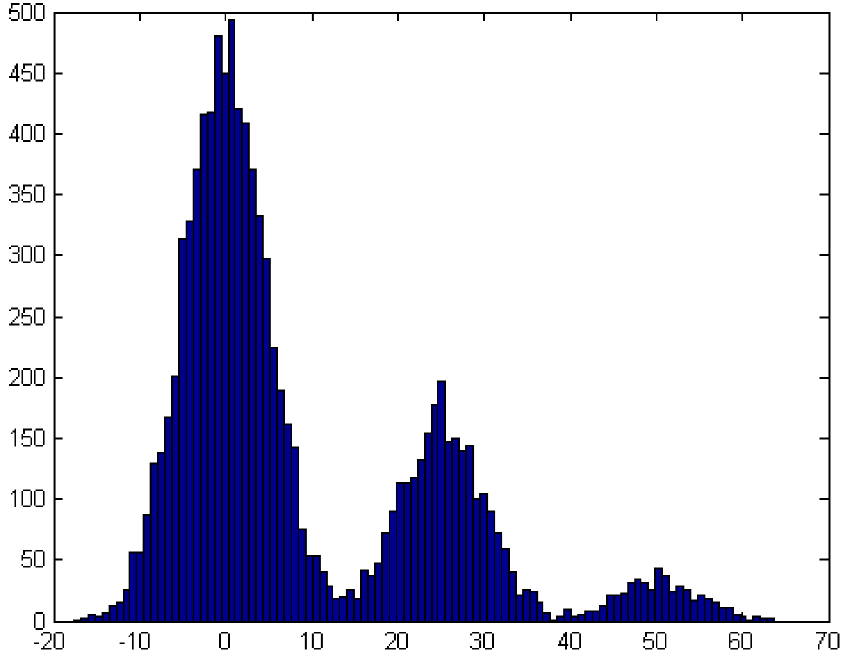
```
a1 = sum(O1SL == 0)/length(variable);
a2 = sum(O1SL == 25)/length(variable);
a3 = sum(O1SL == 50)/length(variable);
a4 = sum(O1SL == 75)/length(variable);
a5 = sum(O1SL == 100)/length(variable);

v1 = var(variable)^0.5 /5;
m1 = 0;
m2 = 25;
m3 = 50;
m4 = 75;
m5 = 100;

Y1 = normrnd(m1,v1,[10000*a1,1]);
Y2 = normrnd(m2,v1,[10000*a2,1]);
Y3 = normrnd(m3,v1,[10000*a3,1]);
Y4 = normrnd(m4,v1,[10000*a4,1]);
Y5 = normrnd(m5,v1,[10000*a5,1]);

Y = [Y1;Y2;Y3;Y4;Y5];
figure
hist(Y,100)

for i=1:1000
Yr = Y(randperm(length(Y)));
Yr1 = Yr((Yr>=0) & (Yr <= 100));
Yr20 = Yr1(1:20);
if( (mean(Yr20) > abaixo do valor da média) && (mean(Yr20) < acima do valor da média))
    if( (var(Yr20) > abaixo do valor da variância) && (var(Yr20) < acima do valor da variância) )
        if( (var(Yr20) > abaixo do valor da desvio) && (var(Yr20) < acima do valor do desvio) )
  break;
end
end
end
```

2 – Result of the Matlab Script






```
. polychoric MISS_PRE_SEQL MISS_POS_SEQL

Variables :  MISS_PRE_SEQL MISS_POS_SEQL
Type :       polyserial
Rho          = .96090913
S.e.         = .01914865
```




```
. polychoric MISS_PRE_SIML MISS_POS_SIML

Variables :  MISS_PRE_SIML MISS_POS_SIML
Type :       polyserial
Rho          = .98219031
S.e.         = .01779227
```