Spot the Difference:

Similarity Within and Across Multiple-suspect and Single-suspect Lineups

Kate Southwood STHKAT006

Veerle Vijverberg VJVVEE001

Eyewitness and ACSENT Laboratories

Department of Psychology

University of Cape Town

18 November 2021

Supervisor: Colin Tredoux

Word count:

Abstract: 250

Main text: 7,685

Main text including figures: 7,997

PLAGIARISM

DECLARATION

1. I know that plagiarism is wrong. Plagiarism is to use another's work and to pretend that it is one's own.

2. I have used the *American Psychological Association (APA)* convention for citation and referencing. Each significant contribution to, and quotation in, this essay / report / project / from the work, or works, of other people has been attributed, and has cited and referenced.

3. This essay /report /project / is my own work.

4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

5. I acknowledge that copying someone else's assignment or essay, or part of it, is wrong, and declare that this is my own work.

NAME: Kate Southwood **SIGNATURE**: 

NAME: Veerle Vijverberg **SIGNATURE:** 

Abstract

There is a lack of research into multiple-suspect lineups, and a lack of explicit comparison between single- and multiple-suspect models, which leads to the question of which model accommodates the highest rate of correct identifications. There is also a question of similarity levels in lineups, across both perpetrator-suspect similarity and lineup member similarity, and how these similarity levels may moderate potential differences in discriminability between the lineup models. To test these questions, a 2x2x2x2 between-subjects experiment was conducted. The variables manipulated were lineup model type (single-suspect or multiple-suspect), similarity level between lineup members (high or low), similarity level between the perpetrator and the suspect (high or low), and target presence (present or absent). The experiment ran as an online survey. Data was analysed using descriptive analysis, logistic regression, and factor analysis. The number of participants in this study was 465 (n = 465). The multiple-suspect lineup model resulted in greater proportions of accurate identifications than the single-suspect model. Single-suspect lineups were found to pose higher risk of making false rejections and foil identifications than the multiple-suspect counterpart. No interaction effects were found between either type of similarity and the lineup model type, however low similarity lineups tended to yield more accurate results. High levels of suspect-perpetrator similarity were found to increase the sense of task difficulty in making lineup decisions. Results prompt further investigation into the potential moderating effect of similarity levels, both of lineup members, and between suspects and perpetrators, on the relationship between lineup model type and accuracy.

Keywords: eyewitness, identification, lineups, multiple-suspects, similarity

Spot the Difference: Similarity Within and Across Multiple-suspect and Single-suspect Lineups

Lineups are a key tool in legal systems, whereby an eyewitness identifies the individual(s) who they believe to be the perpetrator of a crime out of a group of similar-looking lineup members. There are several models of lineups, with debate around which is superior. The consensus in the research community is that the superior model has the highest probability of a correct identification, weighed against the chance of a false identification. While there is research comparing a so-called 'all-suspect' lineup model (every member of the lineup is a suspect) with a 'single-suspect' lineup model (only one member of the lineup is a suspect), there is no substantial research on an interesting alternative known as the 'multiple-suspect model' (more than one suspect in the lineup, but not every person is a suspect). The law requires lineup members to be physically similar to one another, yet this level of similarity is only vaguely defined. Under both signal detection theory and diagnostic feature-detection theory, lower levels of similarity have been argued to improve the chance of correctly identifying perpetrators. It is thus vital to consider similarity, between both the suspect and the perpetrator, as well as between lineup members, as potential moderating factors in any advantage of one lineup model over another.

Background on Lineups

Types of lineups and how they are compared

Before evaluating the lineup types, key terms in the field should be defined. The person who committed the crime is the 'perpetrator', while the person who witnessed this event is the 'eyewitness'. A 'suspect' is the person who the police, after investigating, believe could be the perpetrator. If the eyewitness identifies the suspect as the perpetrator, they are processed through the legal system to be charged for the crime. The following definitions are laid out as outlined by Wells and Turtle (1986), whose research was used as a foundation for

US lineup policy. The chief goal of a lineup is to identify a suspect in a crime. It is important for identifications to be correct, rather than false identifications (identifying an innocent suspect), or foil identifications (identifying a foil) (Wells & Lindsay, 1980). Foils are members of the lineup who are known to be innocent. If they are identified, there are no legal repercussions for that individual.

Assuming that a suspect is in the lineup, three types of lineup models are recognised by Wells and Turtle (1986): a single-suspect lineup consisting of only one suspect, with the remaining members of the lineup being foils, an all-suspect lineup where every member of the lineup is a suspect, and a multiple-suspect model. The multiple-suspect model is a combination of these two models, with more than one suspect in the lineup. The remaining members are foils, whose number increases as the number of suspects increases. All lineups have the legal requirement that members need to be physically similar to one another, although the level of similarity is undefined (Nortje et al., 2020; Wells & Turtle, 1986). Lineup models are often compared using a diagnosticity ratio: the ratio of the probability of a correct identification to that of a false identification (Wixted & Mickes, 2014). All-suspect lineups have low diagnosticity ratios, as the chance of correct identification is high (due to the increased number of suspects), yet the chance of a false identification is also high (Wells & Turtle, 1986). The single-suspect model has a lower chance of correct identification, but an even lower chance of false identification due the use of foils - and thus a higher overall diagnosticity ratio - so the single-suspect model is deemed superior (Wells & Turtle, 1986). However, there is no substantial research on the multiple-suspect model (Nortje & Tredoux, 2021). This model reduces the chance of false identifications through the presence of foils, while increasing the probability of the perpetrator being present by including more than one suspect. Consequently, the multiple-suspect model may have a higher level of diagnosticity

than the single-suspect model, and should thus be considered within the debate for the optimal lineup type.

Current state of lineups in SA

According to the official South African Police Services (SAPS) (2007) identification parade guidelines, lineups with multiple suspects who look dissimilar should either have enough foils to resemble all the suspects (with the number of foils increasing proportionately) or should be held as separate single-suspect lineups. From legal and research commentaries respectively, both du Toit et al. (2019) and Wells et al. (2020) made recommendations to only include multiple suspects in the same lineup if they look sufficiently similar. Live lineups are currently mandatory in South Africa (Fitzgerald et al., 2018), and with single-suspect lineups being the preference, this brings many practical complications for the police, courts, and researchers: the easiest foils to use are prisoners (Nortje, 2018), so lineups are expensive and require extra security and time to organise. Under these current guidelines, if there are multiple suspects in the case (which is the norm), it becomes increasingly difficult to find enough suitable foils for the lineups to work. Thus, there is a recognition within SA to investigate multiple-suspect models as an official lineup option, as it may be more effective to correctly identify suspects, as well as be practically easier to administer.

Testing for Similarity

Why test for similarity?

Similarity between lineup members is required to maintain lineup fairness (Colloff et al., 2021; Malpass et al., 2007; McQuiston-Surrett et al., 2006). There are various methods to select foils, with matching the perpetrator's description being preferred, as this is associated with a lower false-alarm rate (Colloff et al., 2021; Fitzgerald et al., 2013; Luus & Wells, 1991; Wells et al., 1993). In high-similarity models, the eyewitness may struggle to discriminate between lineup members, lowering the accuracy of their identification (Bergold

& Heaton, 2018; Fitzgerald et al., 2015; Malpass et al., 2007; Smith et al., 2018). Yet, moderate-similarity lineups are associated with a higher rate of false identifications than high-similarity lineups (Fitzgerald et al., 2013). Comparatively, low-similarity lineups are linked to more accurate identifications and rejections than moderate- or high-similarity lineups (Bergold & Heaton, 2018; Fitzgerald et al., 2013, 2015; Malpass et al., 2007). So, in order to reduce the rate of misidentification, it may be desirable to conduct a multiple-suspect lineup with a low-similarity level. However, if a suspect is too distinct from other lineup members, the lineup is considered biased (Fitzgerald et al., 2013), as the eyewitness may simply identify an individual as the perpetrator because they stand out (Malpass et al., 2007). When comparing lineup models, any conclusions drawn about one model over another will thus have to consider the role of the level of similarity in moderating its advantage.

How to test for similarity

Researchers in the eyewitness field have used many techniques to measure facial similarity, the most common involving ratings of independent participant judges (Fitzgerald et al., 2013, 2015; Flowe & Ebbesen, 2007; Flowe et al., 2014; Tredoux, 2002). Researchers have also adopted scaling techniques that incorporate such ratings into analytic models measuring spatial distance, which allows for comparison across larger sample sizes (Fitzgerald et al., 2015; Gepshtein et al., 2020; Hirschberg et al., 1978; Tredoux, 2002).

Recent work in the area suggests two theories to explain how the level of similarity plays a role in lineups, namely signal detection theory and diagnostic-feature-detection theory. Signal detection theory is a computational framework that explains how to distinguish a stimulus (signal) from potential distractors (noise) (Stanislaw & Todorov, 1999). It offers a useful way of conceptualizing eyewitness identification, as it differentiates discriminability (the ability to distinguish between perpetrators and innocent suspects and foils) from response

bias (the tendency to identify someone even if unsure of the lineup member's guilt) (Wixted & Mickes, 2014).

The second framework is diagnostic-feature-detection theory, which outlines some features of lineup members as diagnostic (different for innocent and guilty suspects), while others are non-diagnostic (the same for innocent and guilty suspects) (Carlson et al., 2021; Wixted & Mickes, 2014). For example, if the perpetrator was a white woman, all members of the lineup should have the same non-diagnostic features (such as race, build, age), so that the eyewitness only has to differentiate between diagnostic features (such as specific facial features) (Wixted & Mickes, 2014). Higher similarity between lineup members means that fewer features will be diagnostic. This implies that having low-similarity lineups may be more effective in eliciting accurate eyewitness identifications. However, as previously outlined, too little similarity makes the lineup biased. This leaves the question of what the optimal level of similarity in a lineup should be.

Thus, both signal detection theory and diagnostic-feature-detection theory illustrate that differing levels of similarity can influence the eyewitness's ability to discriminate within the lineup models, and should be investigated.

Research Aim and Question

Considering what is currently known about single- and multiple-suspect lineups, two research questions emerge: the first is whether or not single-suspect lineups are superior to multiple-suspect lineups. This question arises from the overall lack of research into multiple-suspect lineups, and lack of explicit comparison between the two models. The second research question regards the issue of similarity levels in a lineup. This relates to two types of similarity: the similarity level between suspects and the perpetrator, and the similarity level between the lineup members. The question is whether these similarity levels moderate potential differences in discriminability and diagnosticity ratios between the lineup models.

The search for the optimal lineup model continues, and these research questions prompt investigation into where multiple-suspect models fit within this field, and at which level of similarity.

Based on existing research, the multiple-suspect model is hypothesised to yield higher accuracy ratings than the single-suspect lineup model conditions overall. Regarding the level of similarity, the lower levels of similarity should yield higher accuracy. Thus, it follows that the multiple-suspect model with low similarity levels will yield the highest accuracy and lowest inaccuracy scores. Even in low similarity conditions, a baseline level of similarity should be upheld to avoid creating biased lineups. The final question being investigated, which it is unclear what to expect, is whether the level of similarity will moderate the advantage that one lineup model type may have over another.

Method

Design

The experiment had a 2 (Type of Lineup: Single-Suspect vs. Multiple-Suspect Lineup) x2 (Target Presence: Target-Present vs. Target-Absent) x2 (Lineup Member Similarity: High vs. Low) x2 (Suspect-Perpetrator Similarity: High vs. Low) between-subjects design, with lineup model type and the respective similarity levels (suspect-vs-perpetrator and lineup member similarity) as independent variables, and accuracy of eyewitness identification as the dependent variable. Data was collected via an online survey. Participants were deceived about the nature of the study, believing they were being surveyed about perceptions of crime. This ensured realistic encoding of the scenario, as witnesses in real life do not expect to see a crime before it happens. Participants watched a video of a crime, of which there were two versions, to account for a possible stimulus effect (further outlined in the *encoding* section). They were then asked to identify the perpetrator from a lineup condition that they were randomly allocated into. A distractor task was conducted

between these two steps, to ensure the perpetrator's face was not still in the participants' working memory. The lineups were constructed using a separate sample of participants, recruited through convenience sampling, who selected lineup images from a larger pool provided by the researchers. As there were several types of analyses to be conducted, the smallest required sample size is 448, which achieves a power of 0.8 in the logistic regression analysis, and more than 0.8 in the other analyses proposed. We arrived at this number after computing the main effects of a simulated version of this study at a power of 0.8, which calculated a sample size 222. Since interaction effects were also to be investigated, we doubled the sample size (Wahlsten, 1991), then rounded up to the nearest number divisible by 16 (ensuring a consistent number of participants in each condition). See Appendix A for power and sample size calculations.

Lineup construction

Participants

Participants for this phase of the study were recruited through convenience sampling. This method was used as the lineup construction was taking place during the UCT vacation period, where access to pools of student participants was limited. See a copy of the consent form at Appendix B.

Materials

Image pool

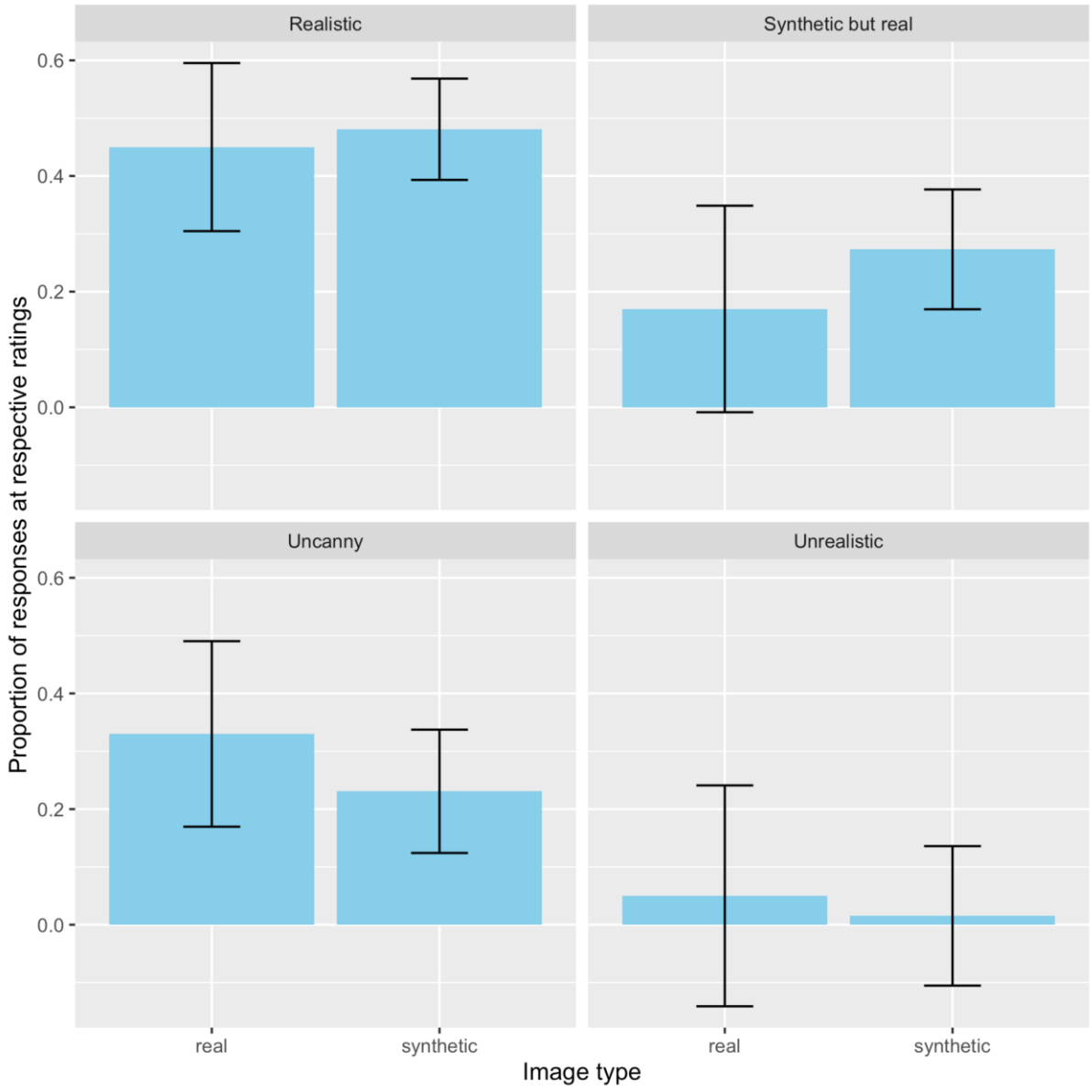
The researchers created a pool of 80 images of possible lineup members, with varying levels of similarity to the perpetrator. These images were synthetic, rather than photographs of real people from a database, to allow for the most accurate manipulation of similarity levels. Refer to Appendix C to see the image pools and selection sheets. The images were created through a combination of a morphing programme called *Facemorph* (accessible at www.facemorph.me) and a synthetic face generator called *Anonymiser* (accessible at

www.generated.photos/anonymiser.com). The images were labelled (01-80), and compared to a reference photograph of the perpetrator and, once selected, the suspects. As there were two versions of the video, and thus, two different perpetrators, this process was completed twice, with the image pools being created based on each possible perpetrator.

A small sample (n = 20) was recruited to rate the realism of these synthetic images. An image pool was created specifically for this test, featuring several real faces and several synthetic images, most of which were morphed images of real and generated faces. Participants were told that some images were real, and some were synthetic, and asked to rate them on a scale of 1 to 4, with 1 being “unrealistic”, 2 being “synthetic but realistic”, 3 being “realistic” and 4 being “uncanny” (which means an image is very realistic, but something about it feels artificial (Seyama & Nagayama, 2007)). Figure 1 is a bar graph depicting the proportionality of these realism ratings. Post-hoc analysis of these results indicates there was not a huge difference in the ratings of real and synthetic images, with close proportions for the *realistic* category. While there was a higher proportion of synthetic images rated as *synthetic but realistic*, this does not indicate that participants were able to tell the difference between image types, as a higher proportion of the *uncanny* ratings were made up of the real images. These results indicate that participants could not tell real images apart from synthetic ones, meaning that the stimuli were of a realistic enough quality to use as lineup images.

Figure 1

Proportion of realism ratings across real and synthetic images



Note. Error bars are 95% confidence intervals.

Selection sheet

Participants selected the images they believed most appropriate for each level of similarity. Participants were randomly assigned to an image pool. Refer to Appendix C for image pools and selection sheets.

Procedure

Each participant selected the top ten images from the total pool for each of the levels of similarity, namely those that were of high similarity to the perpetrator, and those that were low similarity to the perpetrator (meaning they selected 20 images out of the pool). The researchers analysed the selection sheets to determine which images were most often selected, and used these to construct the lineups. Since the experiment participants would only see one lineup condition, the same lineup members were used for the single-suspect lineup and multiple-suspect lineup conditions of corresponding similarity levels, which is why the lineups were only being constructed across similarity levels. While effective size would typically be calculated (Malpass, 1981), the manipulation of similarity levels taking place in this experiment leaves effective size meaningless, as the lineups would necessarily have different selection rates of lineup members. Thus, effective size was not calculated in this instance.

Experiment

Participants

Using the UCT Student Research Participation Programme (SRPP), which supplies undergraduate psychology students for research studies, participants were recruited through convenience sampling. Participants were compensated with one SRPP point for their completion of the survey, which is a requirement for many undergraduate psychology courses. See Appendix D for recruitment announcement. As there was a shortage of participants, the study was closed on the SRPP platform and re-opened and distributed across various other social media platforms, with the incentive to participate being entry into a raffle to win a cash prize (of which there was one valued at R500 and two valued at R250 each). The recruitment announcement for this stage of the study can be found as Appendix E.

Inclusion Criteria

Participants were required to have access to a desktop computer or laptop, as the survey was designed with these devices in mind. Internet connectivity was required for the duration of the experiment. Participants had to be over the age of eighteen years. Participants were not to have participated in an eyewitness identification study in the past eighteen months, although this criterion was not expressed on the consent form as they were unaware that they were going to be asked to make a lineup identification. Instead, the exclusion criterion was phrased as "has not participated in a study relating to crime", as most identification studies are related to witnessing a crime and making an identification. Participants were not excluded on the grounds of eyesight issues, however they were warned that they were going to watch a video.

Materials

Simulated crime video

Participants watched a video of a simulated crime specially scripted and filmed for this study by the authors. In the video, the perpetrator is walking down a street, looking into various gardens. As the perpetrator is outside, the victim walks out the gate of his home, and looks around (as if looking for his lift). He puts his bag down and steps away, distractedly checking his phone (we assume he is contacting his lift). The perpetrator looks around, realizes he has an opportunity, and grabs the bag and runs off. The victim notices the perpetrator, running after him, after which the video ends. There were two versions of the video, with the only difference being the switching of the actors' roles, to account for any stimulus effect that may have been present (such as one actor being too unique-looking that they were more identifiable than average, which could have skewed results). Participants were randomly allocated into viewing either one of the videos. Each video is approximately 40 seconds in duration. The consent form for the actors is attached as Appendix F.

Lineups

There were sixteen conditions which each include different lineups (see Table 1, which is further explained under *lineup selection*). The process of how these lineups were constructed is outlined in the previous section of this proposal. Note that these conditions exist for each video stimulus (so there were two versions of Table 1, with the same layout, but built around different perpetrators and suspects).

Procedure

The experiment was conducted on the online survey platform *Qualtrics*. Participants were randomly allocated to view one of two versions of a video of a crime, completed a distractor task, and were then randomly allocated into a condition with a specific lineup model and level of similarity (between the suspects and perpetrator, and between the lineup members, respectively; see Table 1). They were then asked to select which member of the lineup they believed to be the perpetrator (but had the option to reject the lineup if they did not believe the perpetrator was present). Participants in the single-suspect lineup conditions made their final decision after viewing both lineups, but were able to make a preliminary identification after the first lineup. The participants then completed a questionnaire about their decision-making process, using Likert scale responses. This data was collected because there is evidence that decision latency can be indicative of accuracy of identification (Wittwer et al., in press). The items for this questionnaire have been adapted from Wittwer et al. (in press), included in Appendix G.

Encoding

The participants were told that they were going to watch a video, to ensure they were paying attention to it. However, they were not informed that the video is of a crime, nor that they were going to make an identification, to ensure incidental encoding (as witnesses to crimes are not aware they are going to be witnessing the crime beforehand).

Distractor Task

To ensure the participants have focused on something else so the trace of the perpetrator's face is no longer in their working memory, participants completed two short distractor tasks, which included a flag naming task and a language unscrambling task. They were not aware the next step would be a lineup.

Lineup Selection

At this point, all participants had seen one of the two videos and completed the same distractor tasks. Participants were then randomly allocated into one of the experimental conditions (see Table 1). These are first split according to the level of similarity between the suspects and the perpetrator, this being either low or high. Further, participants would either see two single-suspect lineups (each with one suspect and five foils) or a multiple-suspect lineup (with both suspects in the same lineup, accompanied by ten foils). With regards to the level of similarity between the lineup members, no lineup member should stand out from the others, and the level of similarity should be consistent between them: either all very similar, or only slightly similar (high vs low). Finally, the participants would view a lineup where the target would either be present (TP), meaning the guilty suspect is in the lineup, or they would be absent (TA), meaning the only correct response is to reject the lineup. The TA conditions exist to account for "noise", ensuring participants were not making identifications when they believe the guilty suspect is not present, for the sake of deciding (otherwise known as response bias).

Table 1

Experimental design

Lineup model type	Similarity between suspect and perpetrator			
	Low		High	
	Lineup member similarity[a]		Lineup member similarity[a]	
SSL[b]	Low	High	Low	High
	TA_1-TA_2 Or TA_2-TA_1	TA_1-TA_2 Or TA_2-TA_1	TA_1-TA_2 Or TA_2-TA_1	TA_1-TA_2 Or TA_2-TA_1
	TP-TA Or TA-TP	TP-TA Or TA-TP	TP-TA Or TA-TP	TP-TA Or TA-TP
MSL	TA TP	TA TP	TA TP	TA TP

Abbreviations. TP refers to target-present lineups, where the perpetrator is present. TA refers to target-absent lineups, where the suspects are innocent (not the perpetrator). SSL refers to single-suspect lineups, with one suspect and several foils. In this case, there were two lineups in each SSL condition. MSL refers to a multiple-suspect lineup, with more than one suspect in a lineup, along with a suitable number of foils for each suspect.

[a] 'Lineup member similarity' refers to the similarity between the suspects and the foils, as well as the similarity between the foils themselves. In the case of multiple-suspect lineups

where there is more than one suspect in one lineup, this refers to the lineup members, as well as the similarity between the two suspects.

[b] In the single-suspect TA-TA condition, the order of which lineup was presented first was counterbalanced to ensure a 50-50 chance of seeing Suspect A first and Suspect B second, or Suspect B first and Suspect A second. This is indicated by subscript lettering. In the single-suspect one-innocent-one-guilty lineup (TA-TP or TP-TA) options, the order of lineups presented was also counterbalanced to ensure a 50-50 chance of seeing the guilty suspect lineup first and the innocent suspect lineup second, or vice versa.

Metacognition questionnaire

Participants were asked 15 short questions after they had made their identification. The questions covered their confidence in their choices, and how they came to their decisions (Appendix G). This is theorised to contribute towards accuracy of identification (Wittwer et al., in press), and may differ across lineup models, and is thus important to know. The responses were recorded in Likert-scale responses, so that the data can be analysed through statistical means such as a factor analysis.

Data Analysis

Analysis of data was done using the statistical programming software *R Studio* (RStudio Team, 2021), and R (R Core Team, 2021), which underlies it. Alpha level was set at .05, as is standard practice (Lachenicht, 2017). Descriptive analysis of the main effects was conducted, with the binary outcome variable of accuracy. Further descriptive analysis was conducted with a multinomial outcome of identification type, which was investigated in relation to the lineup model main effect, and across the test conditions.

A binary logistic regression was conducted, with lineup model (single-suspect vs. multi-suspect), level of lineup similarity (between lineup members), and level of suspect

similarity (between suspect and perpetrator) as the predictor variables, and identification decision (accurate identification vs. inaccurate identification) as the outcome variable.

A multinomial logistic regression was also conducted, with lineup model (single-suspect vs. multi-suspect), level of similarity (between lineup members), and level of similarity (between suspect and perpetrator) as the predictor variables, and detailed identification decision (correct identification vs. correct rejection vs. foil identification vs. false identification vs. false rejection) as the outcome variable.

Further, a factor analysis was conducted on metacognition questionnaire responses. The choice to do statistical analysis on this data rather than collect qualitative data and do a thematic analysis is due to time constraints for the researchers.

Ethics considerations

The only potential risk identified by the researchers is that the crime video may have been triggering to participants who have been victims of crimes. However, to account for this, the simulated crime was non-violent. Further, the contact information for psychological support was provided in the consent form, should the participants need it. The debriefing form explicitly stated that the crime witnessed was simulated. Participants were compensated with one SRPP point, or with entry into a raffle to win a voucher (depending on the stage of the study when they participated). Application for ethics approval from the UCT Research Ethics Committee was approved (see Appendix H).

Consent, voluntary participation, and confidentiality

There were two informed consent forms, tailored to the different recruitment stages. A copy of our informed consent form for the SRPP participants is shown in Appendix I. The informed consent form for externally recruited participants is shown in Appendix J. Participation in this study was voluntary, and participants were made aware that they can withdraw their participation at any time without repercussion (although SRPP points or entry

into the raffle were only awarded for full participation). Any information shared by participants would remain confidential. All participants were given an informed consent form which went through the necessary details of this study, although there was some deception, as outlined under *encoding*. However, all deception was minimal, and justified and explained in detail in the debriefing process.

Debriefing

A copy of the debriefing form can be viewed in Appendix K. Once participants had answered the questionnaire, they were thanked for taking part in the study. As the first stage of data collection (SRPP recruitment) occurred through convenience sampling (meaning many participants may interact with one another) over approximately two weeks, we only debriefed participants after the data collection for that phase of the study was complete. This was to prevent potential participants from knowing what the study was about before they participated, to ensure that the results were not contaminated. Participants from the second phase of the data collection (external recruitment) were emailed the debriefing letter within 24 hours of completing their participation. The debriefing letter included an alert to participants that they had the option to request the results of the study once it was complete.

Significance

As outlined in the literature, it is important to know which lineup model type is most accurate. False identifications can lead to innocent suspects being criminalised, and lack of identification leads to perpetrators going free. The multiple-suspect lineup model presents a promising alternative to single-suspect- and all-suspect lineup models. The model appears to maximize the benefit of both of the single-suspect (known error with misidentifications) and all-suspect models (increased chance of target presence), while reducing the associated risk of each. Similarity is legally required for conducting multiple-suspect lineups in South Africa, yet this level of similarity remains undefined. Thus, there exists a gap in the literature

regarding which lineup model is most accurate, and whether the level of similarity between lineup members, as well as between the suspects and perpetrator, moderates any advantage that one model may have over another.

Results

Data collection concluded with 465 participants, with between 28 and 30 participants in each condition. Three types of analysis were conducted. The first was a descriptive analysis, which initially considered the results in terms of accuracy as a binary outcome. An identification decision was regarded as accurate if participants correctly identified the perpetrator in the target-present lineup, or correctly rejected the target-absent lineup. It was regarded as inaccurate if participants identified a foil (foil identification) or another suspect (false identification), or if they falsely rejected a target-present lineup. The results were then considered in terms of the type of identification made, this being one of the five possible outcomes mentioned above. The second type of analysis was an inferential statistical analysis, which included logistic regressions, with binary and multinomial outcome variables. Finally, a factor analysis was conducted to investigate whether lineup types or varying levels of either similarity type being investigated were associated with certain thought processes that accompanied lineup decisions. The results of each of these analyses are presented below. Square brackets represent 95% confidence intervals.

Descriptive analysis

The descriptive analysis provided some clear trends for the data, although not necessarily indicating significance. This analysis was conducted with both binary and multinomial outcomes, which will be presented respectively. The proportion of accurate outcomes across multiple-suspect and single-suspect conditions was 55.13% [46.55%, 63.71%] for multiple-suspect lineup conditions and 44.87% [35.36%, 54.39%] for single-suspect conditions. Across the suspect-perpetrator variable, low levels of suspect-perpetrator

similarity yielded a higher proportion of the accurate outcomes, with 52.99% [44.21%, 61.78%] for the low similarity level and 47.01% [37.68%, 56.34%] for the high similarity level. A similar pattern was seen in the outcomes of the lineup member similarity variable, only to a greater degree, with the low similarity level conditions making up 58.12% [49.83%, 66.41%] of the accurate outcomes, and only 41.88% [32.11%, 51.65%] for the high similarity level conditions. The biggest difference was identified in the target presence variable, with 70.51% [63.56%, 77.47%] of accurate outcomes coming from the target absent conditions, and only 29.49% [18.73%, 40.25%] of the proportion of accurate outcomes coming from the target present lineups, which indicates there were more correct rejections of lineups than correct identifications.

As there were two versions of the video stimulus, they were also assessed based on their accuracy outcomes to check for any stimulus effects which may have been present. Stimulus A made up 48.3% [39.08%, 57.5%] of the proportion of accurate outcomes, and Stimulus B made up 51.71% [42.81%, 60.61%].

Confidence (which was self-scored after making an identification) was analysed in two ways, the first being across the chooser (i.e., made an identification) and non-chooser (i.e., said the perpetrator was not present) groups. It is important to report separately for these groups as confidence has been found to predict accuracy in a more meaningful way amongst those who have made an identification decision than for those who have not (Sporer et al., 1995). See Table 2, which gives a breakdown of these results.

Table 2

Mean confidence scores across choosing groups and accuracy of outcomes

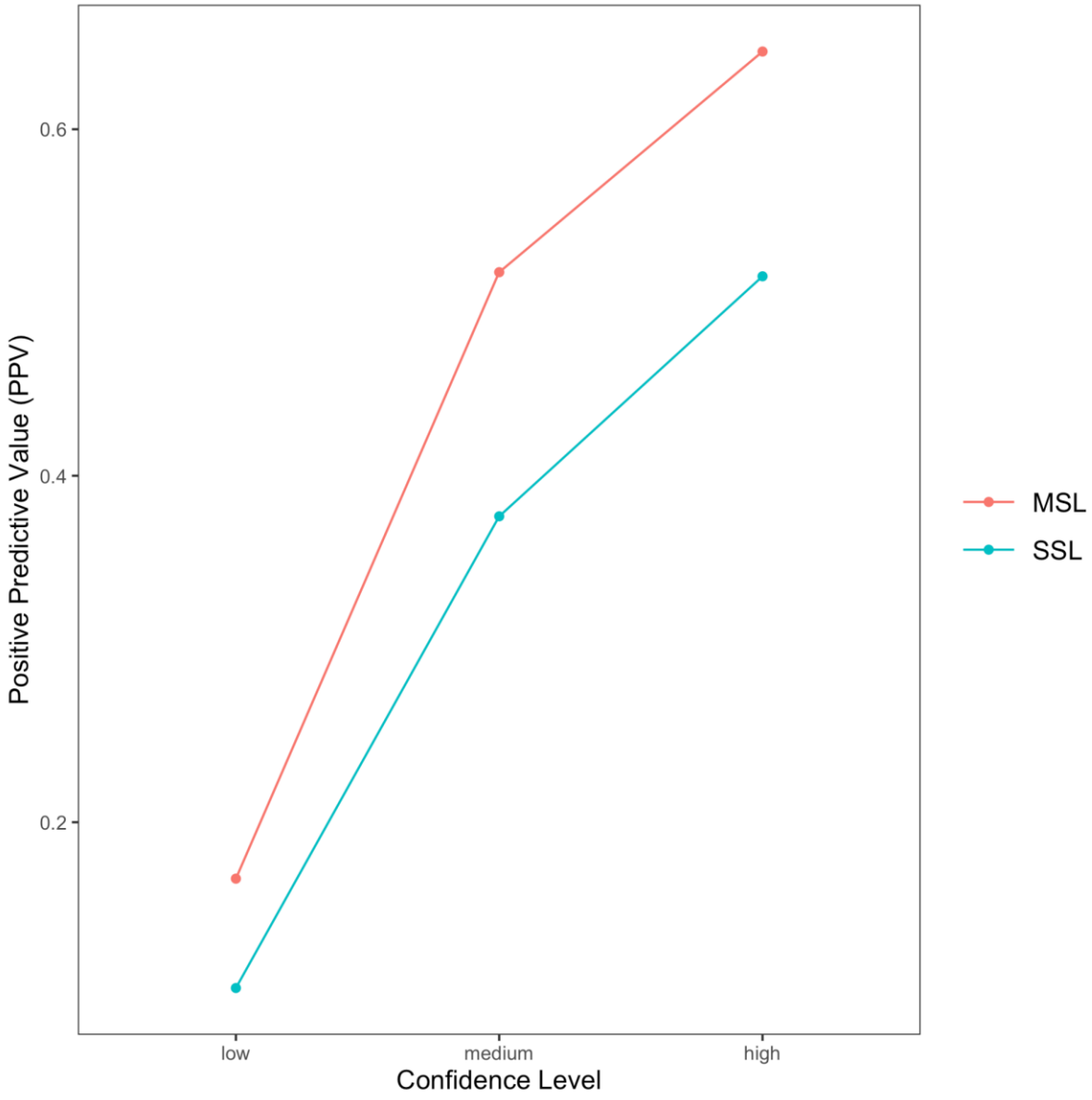
Choosing group	Accuracy	Mean	Standard Deviation
Choosing	Accurate	69.49	21.48
Choosing	Inaccurate	46.33	23.31
Non-choosing	Accurate	63.60	26.05
Non-choosing	Inaccurate	67.13	25.99

Note. All mean and standard deviation values have been rounded to two decimal places.

The group with the highest mean confidence score, and lowest standard deviation, is the choosing group who made accurate identifications (M = 69.5, SD = 21.5). Thus, there is evidence that confidence and accuracy are linked, but only in the choosing group. Thus, confidence was further analysed. See Figure 2, which is a confidence accuracy curve depicting the proportion of accurate responses across confidence groups. This figure indicates a clear pattern, with the highest proportion of those who were accurate being from the multiple-suspect conditions rather than the single-suspect conditions, across all confidence score categories. This means that discriminability (the ability to distinguish between targets that were in the original stimulus from foils and innocent suspects in the lineups) was higher in the multiple-suspect lineups.

Figure 2

Confidence Accuracy Curve



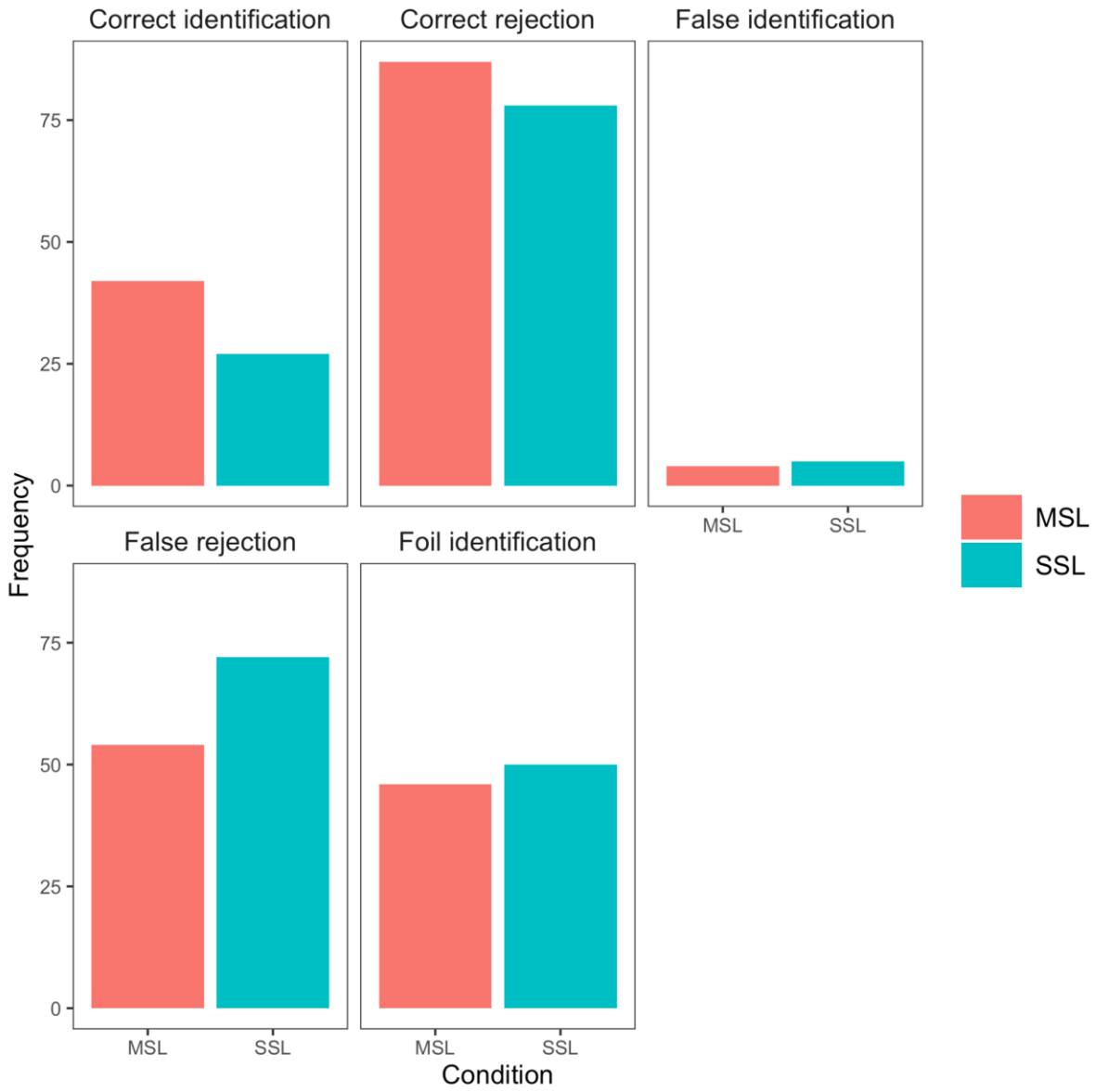
Note. Red line indicates multiple-suspect lineup conditions, blue line indicates single-suspect lineup conditions.

Accuracy was not only considered as a binary outcome, it was further analysed in terms of *type of identification*. Accurate identifications were classified as either correct identifications (in target present lineups) or correct rejections (in target absent lineups). If an eyewitness correctly identifies a suspect as a perpetrator, this means that the perpetrator of

the crime is more likely to face charges for their actions. When a correct rejection is made, this implies that the innocent suspects are allowed to go free. Unfortunately, the police cannot be certain if a correct identification or rejection is made. Types of identification in this study allowed for more extensive coding of decisions: identifications were classified as false identifications (i.e., an innocent suspect was identified), or as foil identifications (i.e., a lineup member who is known to be innocent was identified), or as false rejections (rejecting the lineup when the target was present). If a false identification is made, an innocent suspect is likely to be charged for the crime. This is contrary to a foil identification, where the police know that an eyewitness error has been made, and thus that their identification decision is inaccurate. A false rejection increases the likelihood that a guilty suspect will avoid consequences for their actions. See Figure 3, which is a breakdown of these types of identifications across the test conditions, and notes the difference between the multiple-suspect and single-suspect conditions.

Figure 3

Type of identification across lineup model conditions



Note. Frequency of different identification types across the lineup model conditions, split according to identification types. Note that correct identification and false rejection implies target-present lineup conditions, whereas correct rejection implies target-absent conditions. Foil identifications and false identifications were possible under both lineup types.

Each condition had between 28 and 30 participants. These participants each made an identification or rejected the lineup. Each of these types of identifications gives insights into

the multiple-suspect and single-suspect main effect. Proportionately more accurate identifications were made in the multiple-suspect conditions, making up 60.87% of the correct identifications and 52.73% of the correct rejections. Proportionately less of the inaccurate responses were also made in the multiple-suspect lineup conditions, making up 44.44% of the false identifications, 47.92% of the foil identifications, and 42.86% of the false rejections. The diagnosticity ratio, which is the chance of a correct identification divided by the chance of a false identification, could be calculated for this study. Diagnosticity ratios are a useful calculation to determine whether the trade-off of risking a false alarm for a correct identification is worth it, with higher diagnosticity ratios being prioritised (Wixted & Mickes, 2014). Higher diagnosticity ratios mean better discriminability, and reduced response bias (Wixted & Mickes, 2014). The diagnosticity ratio for multiple-suspect lineup conditions was 1.28 (0.61/0.48), which was higher than the diagnosticity ratio for the single-suspect lineup conditions, which was 0.68 (0.39/0.57). This means that hits (i.e., correct identifications) were 1.28 more likely than false alarms (i.e., foil or false identifications) in multiple-suspect lineup conditions, while hits were 0.68 less likely than false alarms in single-suspect lineups.

Inferential statistical analysis

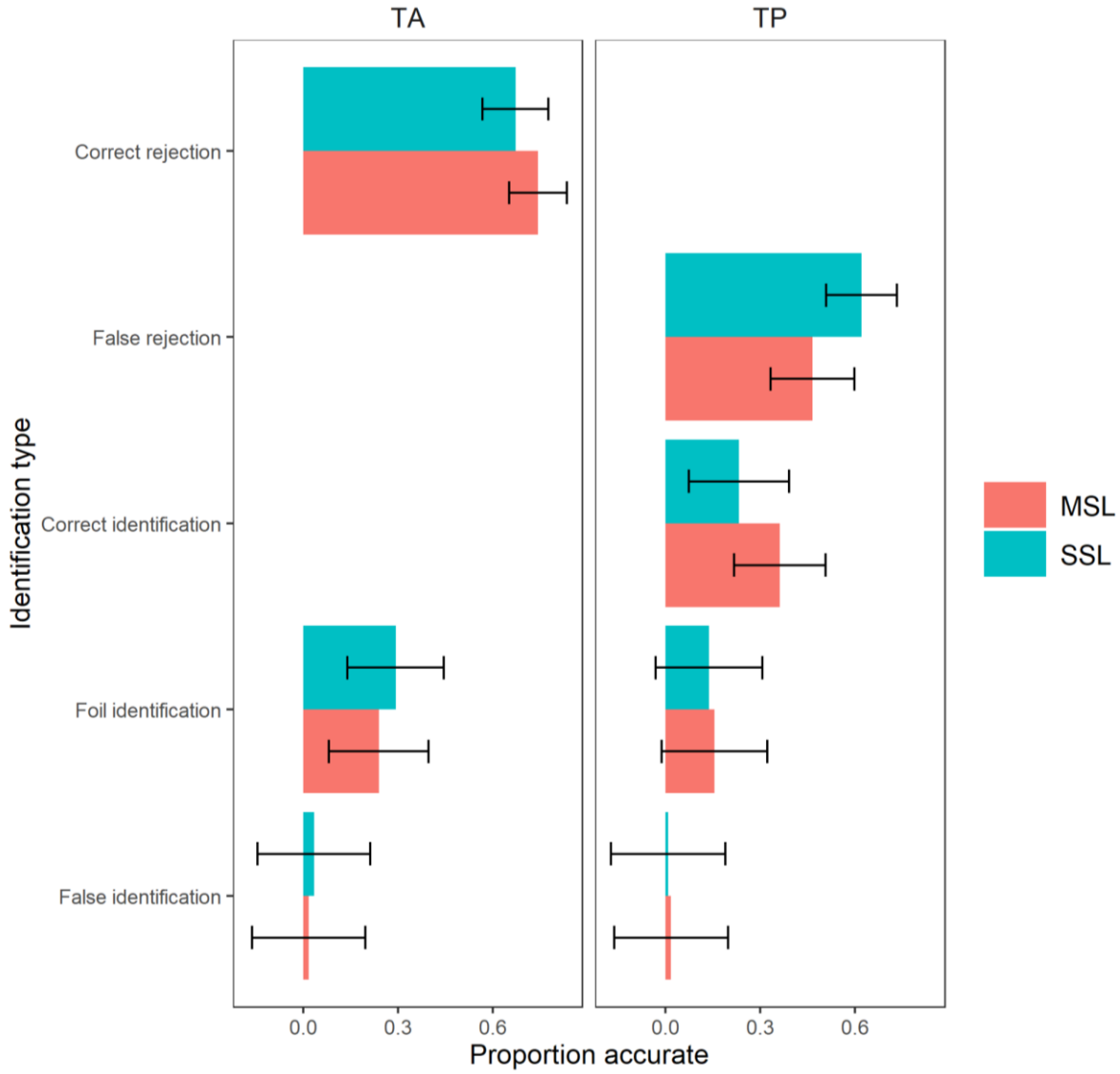
A binary logistic regression was performed to assess the likelihood of lineup type, suspect-perpetrator-, lineup member-similarity, and target-presence affecting the accuracy of lineup decisions. See Table 3 (Appendix L) for a depiction of the results. Here, lineup type (MSL vs SSL) was found to significantly predict accuracy, OR=0.29, z=*-2.22*, p<.03, df=1, SE=0.56, [.09, 0.84], meaning that the odds of making an accurate identification was 29% higher in the multiple-suspect lineup condition. No significant interaction effects were found between lineup type and across similarity levels. There was a non-significant trend for the interactive effect of lineup type and suspect-perpetrator similarity, OR= 1.02, z=*1.75*, p<.08,

df=1, *SE=0.81* [0.85, 20.80]. Post-hoc Tukey tests revealed a statistically significant difference in accuracy between lineup types, z=2.18, p<.03, *SE*=0.19.

Further, a multinomial logistic regression was conducted to consider the type of identification made as a function of lineup type, suspect-perpetrator similarity, lineup member similarity, and target presence. Here, single-suspect lineups were found to associate significantly with an increase in false rejections, OR=0.74, z=-2.22, p=0.03, *SE*= 0.56 [0.12, 1.34]. Hence, the odds of making a false rejection were 26% lower in the multiple-suspect condition. Further, low lineup member similarity was significantly associated with a decrease in the risk of making both foil identifications and false rejections, with an odds ratio of -1.07 [-1.92, -0.22] and -0.64 [-1.24, -.03] respectively. Keeping all other variables constant, an eyewitness is 35% more likely to make a false rejection if the level of lineup member similarity increases by one unit, and 47% more likely to make a foil identification. This means that the risk of making a false rejection is 65% lower in the multiple-suspect condition, and the risk of making a foil identification is 53% lower. Refer to Table 4 (Appendix M) for a depiction of the model coefficients. No interaction effects were observed between lineup types across the similarity levels. Post-hoc tests revealed that while there is a significant difference between correct identifications and foil identifications in the single-suspect condition, this is not the case for the multiple-suspect condition. See Figure 4 for a visual depiction of model results.

Figure 4

Predicted probabilities of accuracy



Note. Relationship between identification type and proportion of accurate decisions made across lineup types, faceted according to target presence. Spaces indicate zero-values, meaning that a specific identification type could not be made for that condition. Error bars are 95% confidence intervals.

Factor analysis

Finally, a factor analysis of the 15-item questionnaire was done to confirm decision strategies previously identified by Wittwer et al. (in press), and to check whether there were

differences in metacognition across the lineup types. A four-factor model was deemed most appropriate, with the factors being named *automatic recognition*, *elimination process*, *sense of familiarity*, and *sense of task difficulty*. See Appendix N, which is a table indicating the loadings onto each of the factors. No significant differences in metacognition were observed between the decisions made across the lineup types. However, a low level of suspect-perpetrator similarity was found to significantly predict the sense of difficulty in making lineup decisions, t=2.12, p=.03, SE= .34 [.05, 1.39]. This means that the sense of task difficulty rose as the level of similarity between the suspect and perpetrator increased. Further, the level of suspect-perpetrator similarity was also found to associate significantly with the adoption of familiarity-search strategies, MSE=3.13, F=4.95 p=.03. In target-present single-suspect lineups, combining low suspect-perpetrator similarity and low lineup member similarity with the adoption of elimination strategies has been found to predict the accuracy of identification decisions, z=-2.09, p=.04, df=1, SE=1.38 [-5.62, -0.20]. The same condition follows a non-significant trend when adopting automatic strategies, z=1.79, p=.07, df=1, SE=2.27 [-0.47, 8.74].

Discussion

There are two key points of discussion that emerge from the results of this study. The first is that of the explicit comparison between lineup model types, including discussion related to confidence. The second is the discussion around the levels of similarity, including both suspect-perpetrator similarity and lineup member similarity.

Lineup Model Types

The data supports the main effect hypothesis that multiple-suspect lineups overall yield higher accuracy scores than single-suspect lineups. This is evident both in the descriptive and statistical analysis outlined above. The multiple-suspect lineup model conditions consistently resulted in more accurate decisions in target-present and -absent

lineups, as well as fewer inaccurate decisions, across false identifications, foil identifications, and incorrect lineup rejections. These results support the theory that multiple-suspect lineup models have a higher diagnosticity ratio than single-suspect lineups, as calculated above. Overall, the data supports the point made by Nortje and Tredoux (2021) that multiple-suspect lineup models are a promising alternative for identification parades in cases where there is more than one suspect.

As illustrated by the confidence accuracy curve (Figure 2), there is a stronger relationship between confidence and accuracy in multiple-suspect lineup conditions than single-suspect conditions. This indicates a stronger discriminability in the multiple-suspect lineup conditions (Nguyen et al., 2017; Wixted & Mickes, 2014), meaning that witnesses are better at distinguishing between the target and the innocent lineup members (Wixted & Mickes, 2014). This higher discriminability means that suspects identified from multiple-suspect lineups are more likely to be guilty than those identified from single-suspect lineups. There may be some concern over the range of the low positive predictive value rates for the low confidence groups in both lineup model types (meaning low confidence is consistently associated with low proportions of accuracy). However, as pointed out by Nguyen et al. (2017), low confidence witnesses are unlikely to be called upon as the final deciding factor in the legal system, so this potential risk is removed.

It should be noted that identification accuracy and the types of identifications made (and thus, the diagnosticity ratio) fluctuated within the respective lineup model type groups, depending on the level of similarity between the suspects-to-perpetrator and the level of similarity between the lineup members in each condition.

Similarity Levels

Suspect-perpetrator Similarity

Ultimately, we found no significant trends in accuracy across suspect-perpetrator similarity levels and lineup types. However, promising non-significant trends prompt further research on this question. Contrary to multiple-suspect lineups, single-suspect lineups with high levels of similarity between the suspect and perpetrator yielded lower accuracy scores than those with low similarity levels. In the multiple-suspect lineup, no significant differences in accuracy were observed across the similarity levels. This implies that the use of multiple-suspect lineups is of particular benefit when the suspect and perpetrator look extremely similar to one another. If it is the case, similar looking suspects should be placed in the same lineup to encourage accurate lineup decisions. Although the feeling of task difficulty does not necessarily predict accuracy (Wittwer et al., in press), it may relate to the level of subjective stress that an eyewitness experiences, in turn affecting their ability to recall the perpetrator. When the suspect and perpetrator look dissimilar, this significantly predicts the use of familiarity search strategies to come to their final decision. The use of search strategies based on a sense of familiarity has been found to significantly predict accuracy (Wittwer et al., in press). From this, it follows that the inclusion of dissimilar looking suspects will enable more accurate identification decisions, by enabling familiarity-search processes in the eyewitness. However, as previous research has shown (Malpass et al., 2007), structuring the lineup in such a way poses a risk to the liberty of innocent suspects. This is as the eyewitness will be more likely to choose the suspect as the perpetrator, as they are accompanied by foils who can easily be discredited.

Lineup Member Similarity

While a low level of similarity between the lineup members was not associated with significant differences in accuracy amongst the lineup types, this was not the case for conditions with high lineup member similarity. These findings suggest that single-suspect lineups with high lineup member similarity should be utilized with caution, to prevent the

identification of innocent fillers and suspects or the rejection of a lineup where the perpetrator is present. If lineup members look extremely similar, they should rather be placed in a multiple-suspect lineup. If single-suspect lineups are used, the level of similarity between lineup members should be low. Yet even multiple-suspect lineups with low levels of similarity between lineup members are associated with more accurate recall of the perpetrator than those with high lineup member similarity. From this it follows that lineups of either model type with lower levels of similarity will yield more accurate identification decisions. However, conducting a lineup with dissimilar-looking lineup members cannot be considered a fair lineup. In order to prevent biasing the lineup towards the suspect, a baseline of similarity between lineup members should be upheld. In future, studies that examine and define this level further may be of particular importance in this area of interest.

Conclusion

Overall, the multiple-suspect lineup model was the more efficient lineup model in terms of maximizing accurate identification decisions while minimizing inaccurate ones. Although no interaction effects were observed between types of similarity across the lineup models, further investigation into potential moderating effects is recommended, as promising nonsignificant trends were noted in the study, with low levels of similarity between lineup members being associated with improved accuracy. These recommendations are largely informed by the limitations of this study.

Limitations

This study's use of synthetic faces instead of photographs of real faces for the lineups raises concerns around the *uncanny valley* effect (Seyama & Nagayama, 2007), more specifically whether their artificiality affected how well the eyewitness recognises the perpetrator's face amongst the other lineup members. Balas and Pacella (2015), for instance, suggest that it is more difficult to recall artificial faces than real ones. However, the similarity

manipulation was best achieved through use of these synthetic faces, as a broader range of faces with high similarity to the perpetrator could be attained. Further, it is easier to manipulate which facial features to combine, and to what degree, than it would be with the use of real photographs. Additionally, the results of the realism rating survey suggest that people struggle to distinguish between real and synthetic faces. Wells et al. (1993) argue that optimal eyewitness performance occurs when foils are selected to be as low as possible in similarity to the suspect while still matching description. Apart from the concerns of lineup fairness, this point raises the question of whether it is appropriate to build lineups with synthetic faces, where a vast array of extremely similar-looking lineup members can be found. More extensive research of the relationship between the faces' realism and the way they are processed will be a useful venture in determining the appropriateness of building lineups with synthetic faces.

The transitivity of similarity was taken for granted in this study. Transitivity of similarity is the phenomenon that if Stimulus A is of a set similarity to Stimulus B, and Stimulus B has that same set similarity level to Stimulus C, it should follow that Stimulus A is also that set degree of similar to Stimulus C. However, this is not something that can be assumed, and thus needs to be confirmed. This phenomenon would apply to the lineup construction, as this study built lineups and manipulated similarity across several relationships (suspect to suspect, each suspect to perpetrator, each suspect to foils), but it did not confirm that these levels of similarity were transitive across all lineup members. Confirming that similarity across lineup members was transitive would have allowed for a fuller understanding of how similarity moderates the advantage one lineup model type may have over another.

Recommendations

This study prompts further investigation into the potential interaction between similarity levels between lineup members and similarity levels between suspects and perpetrators in a lineup, and the accuracy of eyewitness decisions. While the results showed promising nonsignificant trends in this area, a larger sample size should be recruited for a more powerful test. Further, we endorse Wittwer et al.'s (in press) approach to investigating the metacognition that accompanies eyewitness decisions, and recommend applying this to the research topic at hand. In this case, its application has allowed for a more holistic investigation of the debate between lineup models than what statistical means can provide.

In practice, we suggest the use of multiple-suspect lineups in cases with multiple suspects in a single-perpetrator crime, over the use of several single suspect lineups. The use of multiple suspects in a lineup also has practical benefits, such as a reduced monetary and time-related costs. Further, as the findings of this study indicate, multiple-suspect lineup models have a higher diagnosticity ratio than single-suspect lineups, so the likelihood of correct identification and removing the risk of a false identification is improved. Ultimately, the multiple-suspect lineup model offers a promising alternative to the traditional single-suspect lineup model.

References

Balas, B., & Pacella, J. (2015). Artificial faces are harder to remember. *Computers in human behavior*, *52*, 331-337. https://doi.org/10.1016/j.chb.2015.06.018

Bergold, A. N., & Heaton, P. (2018). Does filler database size influence identification accuracy? *Law and Human Behavior*, *42*(3), 227-243. https://pubmed.ncbi.nlm.nih.gov/29809026/

Carlson, C. A., Hemby, J. A., Wooten, A. R., Jones, A. R., Lockamyeir, R. F., Carlson, M. A., Dias, J. L., & Whittington, J. E. (2021). Testing encoding specificity and the diagnostic feature-detection theory of eyewitness identification, with implications for showups, lineups, and partially disguised perpetrators. *Cognitive Research: Principles and Implications, 6*(1), 1-21. https://doi.org/10.1186/s41235-021-00276-3

Colloff, M. F., Wilson, B. M., Seale-Carlisle, T. M., & Wixted, J. T. (2021). Optimizing the selection of fillers in police lineups. *Proceedings of the National Academy of Sciences, 118*(8). https://doi.org/10.1073/pnas.2017292118

du Toit, E., de Jager, F., Paizes, A., St Quintin Skeen, A., & van der Merwe, S. (2019). *Commentary on the Criminal Procedure Act* (63rd edition). Juta

Fitzgerald, R. J., Oriet, C., & Price, H. L. (2015). Suspect filler similarity in eyewitness lineups: A literature review and a novel methodology. *Law and Human Behavior*, *39*(1), 62-74. https://doi.org/10.1037/lhb0000095

Fitzgerald, R. J., Price, H. L., Oriet, C., & Charman, S. D. (2013). The effect of suspect-filler similarity on eyewitness identification decisions: A meta-analysis. *Psychology, Public Policy, and Law, 19*(2), 151-164. https://doi.org/10.1037/a0030618

Fitzgerald, R. J., Rubinova, E., & Juncu, S. (2018). Eyewitness ID around the World. In A. M. Smith, M. P. Toglia & J. M. Lampinen (Eds.), *Methods, Measures, and Theories in Eyewitness Identification Tasks* (pp. 294-322). Routledge. https://osf.io/kn6r5/

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. https://doi.org/10.1007/s10979-006-9045-9

Flowe, H. D., Klatt, T., & Colloff, M. F. (2014). Selecting fillers on emotional appearance improves lineup identification accuracy. *Law and Human Behavior, 38*(6), 509–519. https://pubmed.ncbi.nlm.nih.gov/25133919/

Gepshtein, S., Wang, Y., He, F., Diep, D., & Albright, T. D. (2020). A perceptual scaling approach to eyewitness identification. *Nature communications*, *11*(1), 1-10. https://doi.org/10.1038/s41467-020-17194-5

Hirschberg, N., Jones, L. E., & Haggerty, M. (1978). What's in a face: Individual differences in face perception. *Journal of Research in Personality, 12*, 488–499. https://doi.org/10.1016/0092-6566(78)90074-0

Lachenicht, L. (2017). Statistical power. In C. Tredoux & K. Durrheim (Eds.), *Numbers, hypotheses & conclusions: A course in statistics for the social sciences* (pp. 230-251). Juta.

Luus, C. E., & Wells, G. L. (1991). Eyewitness identification and the selection of distractors for lineups. *Law and Human Behavior*, *15*(1), 43-57. https://doi.org/10.1007/BF01044829

Malpass, R. S. (1981). Effective size and defendant bias in eyewitness identification lineups. *Law and Human Behavior*, 5(4), 299-309. https://doi.org/10.1007/BF01044945

Malpass, R. S., Tredoux, C. G., & McQuiston-Surrett, D. (2007). Lineup construction and lineup fairness. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The Handbook of Eyewitness Psychology*, Vol. II. Memory for people (pp. 155–178). Lawrence Erlbaum Associates Publishers.

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. https://doi.org/10.1037/1076-8971.12.2.137

Nguyen, T. B., Pezdek, K., & Wixted, J. T. (2017). Evidence for a confidence–accuracy relationship in memory for same- and cross-race faces. *The Quarterly Journal of Experimental Psychology, 70*(12), 2518-2534. https://doi.org/10.1080/17470218.2016.1246578

Nortje, A. (2018). The butcher, the baker, the candlestick maker: Investigating facial recognition for multiple-perpetrator crimes [Doctoral dissertation, University of Cape Town].

Nortje, A., & Tredoux, C. (2021). *Line-up models and multiple perpetrators: A separation of theory and practice.* Unpublished manuscript. Psychology Department, University of Cape Town.

Nortje, A., Tredoux, C. G., & Vredeveldt, A. (2020). Eyewitness identification of multiple perpetrators. *South African Journal of Criminal Justice*, *33*(2), 348-381. Retrieved from https://annelies.vredeveldt.com/wp-content/uploads/2021/02/Nortje-et-al.-2020-Eyewitness-identification-of-multiple-perpetrators.pdf

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

RStudio Team (2021). RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA. http://www.rstudio.com/

Seyama, J., & Nagayama, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence, 16*(4), 337-351. https://doi.org/10.1162/pres.16.4.337

Smith, A. M., Wells, G. L., Smalarz, L., & Lampinen, J. M. (2018). Increasing the similarity of lineup fillers to the suspect improves the applied value of lineups without improving memory performance: Commentary on Colloff, Wade, and Strange (2016). *Psychological Science*, *29*(9), 1548-1551. https://doi.org/10.1177/0956797617698528

South African Police Service. (2007). *National Instruction: identification parades*. Pretoria: Commissioner of the SAPS.

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: a meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, *118*(3), 315-324. https://doi.org/10.1037/0033-2909.118.3.315

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers, 31*(1), 137-149. https://doi.org/10.3758/BF03207704

Tredoux, C. (2002). A direct measure of facial similarity and its relation to human similarity perceptions. *Journal of Experimental Psychology: Applied, 8*(3), 180-193. https://doi.org/10.1037/1076-898X.8.3.180

Wahlsten, D. (1991). Sample size to detect a planned contrast and a one degree-of-freedom interaction effect. *Psychological Bulletin*, *110*(3), 587-595. https://doi.org/10.1037/0033-2909.110.3.587

Wells, G. L., Kovera, M. B., Douglass, A. B., Brewer, N., Meissner, C. A., & Wixted, J. T. (2020). Policy and procedure recommendations for the collection and preservation of eyewitness identification and evidence. *Law and Human Behavior, 44*(1), 3-36. https://pubmed.ncbi.nlm.nih.gov/32027160/

Wells, G. L., & Lindsay, R. C. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin, 88*(3), 776-784. https://doi.org/10.1037/0033-2909.88.3.776

Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distractors for eyewitness lineups. *Journal of Applied Psychology*, 78(5), 835-844. https://doi.org/10.1037/0021-9010.78.5.835

Wells, G. L., & Turtle, J. W. (1986). Eyewitness identification: The importance of lineup models. *Psychological Bulletin, 99*(3), 320-329. https://doi.org/10.1037/0033-2909.99.3.320

Wittwer, T., Tredoux, C. G., Py, J., Nortje, A., Kempen, K., & Launay, C. (in press) Automatic recognition, elimination strategy and familiarity feeling: cognitive processes predict accuracy from lineup identifications. Unpublished manuscript.

Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychological Review*, 121(2), 262-276. https://pubmed.ncbi.nlm.nih.gov/24730600/

Appendices

Appendix A - Power and Sample Size Calculations

In order to compute the required sample size for interpreting effects at a power level of 0.8, data for this study was simulated.

This was done using the number of observations reported in a study by Nortje and Tredoux (2021), as well as a probability of 0.5 (as the variables are binary and there is an equal chance of viewing either of the levels).

The minimum required sample size for interpreting main effects was computed: 222 participants. However, in order to interpret interaction effects accurately, the sample size must be more than double that of main effects (Wahlsten, 1991). Thus, a sample size of approximately 444 participants is required to ensure that interaction effects can be interpreted accurately at a power level of 0.8.

Data Simulation

```
set.seed(5701)

SSLv.MSL <- rbinom(46, 1, 0.5)

TAvs.TP <- rbinom(46, 1, 0.5)

SimilarityLM <- rbinom(46, 1, 0.5)

SimilarityPerpSus <- rbinom(46, 1, 0.5)

AccuracyBinom <- rbinom(46, 1, 0.5)

LineupData <- tibble(SSLv.MSL, TAvs.TP, SimilarityLM, SimilarityPerpSus, AccuracyBinom)
```

```
## # A tibble: 46 x 5
##    SSLv.MSL TAvs.TP SimilarityLM SimilarityPerpSus AccuracyBinom
##    <fct>    <fct>   <fct>        <fct>                      <int>
##  1 MSL      TA      Low          High                           0
##  2 SSL      TA      High         High                           1
```

```
##  3 SSL       TP      High            Low                     0

##  4 MSL       TP      Low             Low                     1

##  5 SSL       TA      High            Low                     1

##  6 MSL       TP      Low             Low                      0

##  7 SSL       TA      Low             Low                     0

##  8 SSL       TP      Low             Low                     1

##  9 SSL       TA      Low             Low                     0

## 10 SSL       TA      High            Low                     0

## # ... with 36 more rows
```

Sample Size for Main Effects

Lineup Model.

Computing the sample size required for the main effect of lineup model and accuracy of identification at a power level of 0.8.

First need to determine p(correct|Single-Suspect Lineup) vs. p(correct|Multiple-Suspect Lineup), and p(Multiple-Suspect Lineup)

```
##  SSLv.MSL            0         1

##         SSL 0.6800000 0.3200000

##         MSL 0.5714286 0.4285714

##  SSLv.MSL  n   percent

##         SSL 25 0.5434783

##       MSL 21 0.4565217

p_load(powerMediation)

powerMediation::SSizeLogisticBin(p1 = .32, p2 = .68, B = .457, alpha = .01, power = .8)

## [1] 88
```

Thus, a sample size of 88 is required to interpret the main effect of lineup model on identification accuracy at a power level of 0.8.

Lineup Member Similarity.

Computing the sample size required for the main effect of Lineup Member Similarity and accuracy of identification at a power level of 0.8.

First need to determine p(correct|Low LM similarity) vs. p(correct|High LM similarity), and p(High LM similarity)

```
LineupData%>%
 tabyl(SimilarityLM, AccuracyBinom)%>%
 adorn_percentages()
##  SimilarityLM         0         1
##          Low 0.6071429 0.3928571
##         High 0.6666667 0.3333333
LineupData%>%
 tabyl(SimilarityLM)
##  SimilarityLM  n   percent
##          Low 28 0.6086957
##         High 18 0.3913043
p_load(powerMediation)
powerMediation::SSizeLogisticBin(p1 = .67, p2 = .39, B = .391, alpha = .01, power = .8)
## [1] 152
```

Thus, a sample size of 152 is required to interpret the main effect of lineup member similarity on identification accuracy at a power level of 0.8.

Suspect-perpetrator Similarity.

Computing the sample size required for the main effect of Similarity between suspect and perpetrator and accuracy of identification at a power level of 0.8.

First need to determine p(correct|Low suspect-perpetrator Similarity) vs. p(correct|High suspect-perpetrator Similarity), and p(High suspect-perpetrator Similarity)

```
LineupData%>%
 tabyl(SimilarityPerpSus, AccuracyBinom)%>%
 adorn_percentages()
##  SimilarityPerpSus         0         1
##                Low 0.6153846 0.3846154
##               High 0.6500000 0.3500000
LineupData%>%
 tabyl(SimilarityPerpSus)
##  SimilarityPerpSus  n   percent
##                Low 26 0.5652174
##               High 20 0.4347826
p_load(powerMediation)
powerMediation::SSizeLogisticBin(p1 = .61, p2 = .38, B = .435, alpha = .01, power = .8)
## [1] 222
```

Thus, a sample size of 222 is required to interpret the main effect of suspect-perpetrator similarity on identification accuracy at a power level of 0.8.

Target Presence.

Computing the sample size required for the main effect of Target presence and accuracy of identification at a power level of 0.8.

First need to determine p(correct|Target Absent) vs. p(correct|Target Present), and p(Target Present)

```
LineupData%>%
 tabyl(TAvs.TP, AccuracyBinom)%>%
 adorn_percentages()
##  TAvs.TP     0         1
##        TA 0.6923077 0.3076923
##        TP 0.5500000 0.4500000
LineupData%>%
 tabyl(TAvs.TP)
##  TAvs.TP  n   percent
##        TA 26 0.5652174
##        TP 20 0.4347826
p_load(powerMediation)
powerMediation::SSizeLogisticBin(p1 = .69, p2 = .31, B = .434, alpha = .01, power = .8)
## [1] 80
```

Thus, a sample size of 80 is required to interpret the main effect of target presence on identification accuracy at a power level of 0.8.

Appendix B - Lineup construction consent form

Study Title: Similarity within and across single- and multiple-suspect lineups

Purpose:

You are invited to construct a pool of lineup members who will be presented to eyewitnesses as part of a lineup experiment at a later stage of this study.

Procedure:

- You will be presented with a pool of images, as well as a reference image
    - Each image is labelled
    - The researchers will explain the various types of similarity to you
    - You will identify the top 20 images for each type of similarity (low and high) out of the pool of images. You will make these identifications by filling in a sheet (provided by the researcher) and documenting the corresponding image label.
- This procedure will take approximately 5-10 minutes.

Exclusion criteria:

- None

Risks and benefits

There are no identified risks or benefits to participating in this study. However, should any risk be identified by a participant, the researchers will make every effort to address this risk.

Confidentiality

Your responses will be recorded and kept entirely confidential. Only the researchers will be aware of your identity. Your responses, along with those of other participants, will be used to

compile a list of lineup members out of this pool of images, thus individual responses will not be included in the study. These responses will only be used for academic purposes.

I have read the above description of the procedure and give my voluntary consent to participate. I acknowledge that my participation in this study is completely voluntary, and that I may withdraw my participation from the study at any time. I understand that the information obtained in this study will remain confidential and internal to the study. I have been assured that any questions or concerns I may have in the future will be answered by the researchers.

Name of Participant Signature

_____ _____

Date

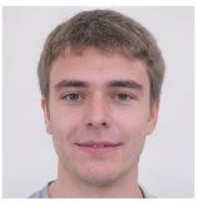
_____

Appendix C - Similarity Ranking and Selection Sheets

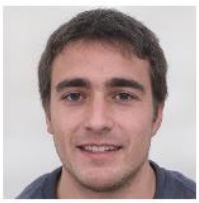
Appendix Ci - Stimulus A image pool

Perpetrator description: blonde hair, blue eyes, oval face shape, facial hair, white skin

				
01	02	03	04	05
				
06	07	08	09	10

				
11	12	13	14	15
				
16	17	18	19	20
				
21	22	23	24	25
				
26	27	28	29	30

				
31	32	33	34	35

				
36	37	38	39	40

				
41	42	43	44	45

				
46	47	48	49	50

				
51	52	53	54	55

				
56	57	58	59	60

				
61	62	63	64	65

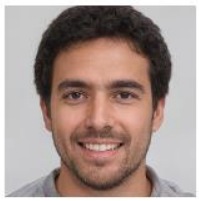
				
66	67	68	69	70

				
71	72	73	74	75
				
76	77	78	79	80

Appendix Cii - Stimulus B image pool

Perpetrator description: Brown curly hair, brown eyes, long face shape, tanned skin, facial hair

				
01	02	03	04	05
				
06	07	08	09	10

				
11	12	13	14	15
				
16	17	18	19	20
				
21	22	23	24	25
				
26	27	28	29	30

				
31	32	33	34	35

				
36	37	38	39	40

				
41	42	43	44	45

				
46	47	48	49	50

				
51	52	53	54	55

				
56	57	58	59	60

				
61	62	63	64	65

				
66	67	68	69	70

				
71	72	73	74	75

				
76	77	78	79	80

Appendix Ciii - Selection sheet

High similarity lineup members:
A lot of the same characteristics to the perpetrator description

Low similarity lineup members:
Only a few of the same characteristics to the perpetrator description

VOTING

Name: _____

You are selecting faces according to similarity (either high or low) to the individual in the large photograph at the top of the first page.
There are no wrong answers.
Please select 20 images that bear high similarity to the individual, and another 20 images that bear low similarity to the individual (40 images in total). Write down the number of the relevant images in the template below.

Images with LOW similarity to individual		**Images with HIGH similarity to individual**	

Thank you for your participation.

Kate Southwood and Veerle Vijverberg
STHKAT006 VJVVEE001

Appendix D - SRPP Recruitment Announcement

Students are invited to participate in an online survey investigating perceptions of crime.

Details about the study: This is an online computer-based survey study. You will be asked to watch a video and answer some questions. You will need to have steady internet access for the duration of this study.

Link to survey: *[will insert link here once survey is live]*

On the completion of your participation in this study, you will receive 1 (one) SRPP point towards your Psychology course requirements.

Inclusion criteria: you must be a UCT student, over the age of 18 years, and not have completed a study relating to crime in the last 18 (eighteen) months.

If there are any issues, or if you have any questions, you can email the researchers directly on:
sthkat006@myuct.ac.za and vjvvee001@myuct.ac.za

Thank you!

Kind regards,
Kate Southwood and Veerle Vijverberg

Appendix E - External Recruitment Announcement

PARTICIPATE IN 15-MINUTE RESEARCH STUDY FOR CHANCE TO WIN R500!

You are invited to participate in an online survey investigating perceptions of crime.

Details about the study: This is an online computer-based survey study. You will be asked to watch a video and answer some questions. You will need to have steady internet access for the duration of this study.

On the completion of your participation in this study, you will receive 1 entry into a raffle, with a chance of winning one of three prizes! One prize is valued at R500, and two prizes are valued at R250 each!

Inclusion criteria: you must be over the age of 18 years, and not have completed a study relating to crime in the last 18 (eighteen) months. You must complete this study on a desktop or laptop computer (as it requires a big screen).

If there are any issues, or if you have any questions, you can email the researchers directly on:
sthkat006@myuct.ac.za and vjvvee001@myuct.ac.za

Thank you!

Kind regards,
Kate Southwood and Veerle Vijverberg

Terms and Conditions:
Raffle winners will be randomly selected by an independent party once data collection for the study is complete. If you are randomly selected to win one of these prizes, you will be contacted via the email address you provided. If you do not reply within one week of being contacted, you will forfeit your prize and another winner will be selected. Note: entry into this raffle does not guarantee winning one of the prizes.

Appendix F - Actor consent form

Study Title: Similarity within and across single- and multiple-suspect lineups

Purpose

You are invited to participate in the filming of a short video which will depict a crime, to be used as encoding material for a larger study on eyewitness identification and similarity and lineup model conditions. The video and accompanying images will also be added to the University of Cape Town Eyewitness Research database, for use in future projects.

Procedure

You will act in two versions of the video, once as the perpetrator of the crime, and once as the victim of the crime.

- The crime will be non-violent, depicting a bag-snatching incident
- The video will be approximately 20 seconds (each version)
- No audio will be used
- Photographs will be taken as well to be used in lineup construction

Exclusion criteria

None

Risk

There is potential that participants of the study may believe the simulated crime is real (and thus, that you committed the crime). However, this will be addressed in the debriefing process where the researchers will alert participants to the fact that the video was a staged crime.

Benefits

There are no personal benefits for participating in the filming of this video.

Confidentiality

Your identity will not be shared at any stage. Only your faces will be used in lineup construction, and in the video, however there will never be any accompanying identifying information with these.

I have read the above description of the procedure and give my voluntary consent to participate, acting in the simulated crime video and being photographed for lineup construction purposes. I acknowledge that my participation in this study is completely voluntary, and that I may withdraw my participation from the study at any time. I understand that other than my face, no other identifying information will be shared, and that there will be a disclaimer that the crime was simulated. Any questions or concerns I have raised have been addressed to my satisfaction by the researchers. I agree that this video and photograph may be added to the UCT Eyewitness Research database, on the condition that any project which uses this video and/or image includes a disclaimer that the crime was staged.

Name of actor _____

Signature of actor _____

Date_____

Appendix G - Metacognition Questionnaire

The following questionnaire has been adapted from Wittwer et al. (in press) to gather metacognition data about the decision-making process involved in lineup identification.

All questions will have a Likert-scale response of:

1 = completely agree

2 = agree

3 = neutral

4 = disagree

5 = completely disagree

1. I immediately recognised the perpetrator, but I cannot explain why.
2. When I viewed the lineup, I immediately recognised the perpetrator because their face jumped out at me.
3. I compared the images with each other to make my choice.
4. I used a process of elimination until I arrived at a final choice.
5. I had no sense of familiarity when I looked at the lineup, but I know I would have recognised the perpetrator if they had been presented.
6. I was afraid of not making the right decision.
7. I tried to remind myself of the perpetrator so that I could compare him to the proposed images.
8. I tried to match the image of the perpetrator that I had in my mind with the images presented.
9. I would have had less difficulty in making my choice if the people had been presented to me in person rather than through images.
10. I immediately recognised someone, then I compared the other images to be sure.
11. I first spotted the photograph that seemed right to me then I looked at the others and gradually eliminated them.
12. I had made a first choice but I eventually chose another person.
13. I imagined each member of the lineup in the theft situation to determine if they were the perpetrator.
14. During the task, I was aware that my memory is not infallible and that I could be mistaken.
15. Some of the images were too similar and it made me doubtful.

Appendix H - Proof of Approved Ethics Application

PSY4025W Ethical Approval Form and Research Proposal

1 message

Kate Southwood <katesouthwood96@gmail.com> Tue, Jun 15, 2021 at 8:58 AM

To: Rosalind Adams <rosalind.adams@uct.ac.za>

Good morning Ros

As instructed, attached is the Ethical Approval Form and Research Proposal for Kate Southwood (STHKAT006) and Veerle Vijverberg (VJVVEE001). Any supporting documentation has been attached as an appendix to the ethical approval form.

Please note Appendix G of the research proposal, as attached to this email, is blank. We will be inserting this email into that appendix as soon as it has been sent. However, the research proposal which we are submitting through Vula will have the updated Appendix G.

Kind regards,
Kate Southwood
Veerle Vijverberg

2 attachments

Kate Southwood STHKAT006 and Veerle Vijverberg VJVVEE001 Ethical Approval Form.pdf
1425K

Research Proposal - Kate Southwood STHKAT006 and Veerle Vijverberg VJVVEE001.docx
691K

Re: Ethics Approval update request

Catherine Ward <cathy.ward.sa@gmail.com> Mon, Sep 6, 9:19 PM

to me, Colin, veerlevijverberg@gmail.com, Veerle, Rosalind

Hi Kate,

Make the one change I've put into your consent form, and your changes are approved.

Best,
Cathy

On Mon, Sep 6, 2021 at 3:33 PM Kate Southwood <katesouthwood96@gmail.com> wrote:

Good day Cathy

I hope this email finds you well.

We have launched our study on the SRPP site, but unfortunately the uptake of participants has been slower than we would like. Since we released the study on Thursday, we have only received 52 participants. As we need a large number of participants (n=672) in an extremely short period of time, we are looking at our other recruitment options.

We have spoken to Colin, and he is prepared to grant us the funds to offer prizes for a raffle as incentive to get participants from other platforms. We understand that

this would deviate from what we have been cleared to do from an ethics perspective.

We have updated our ethics request form, as well as our research proposal. We have made comments or highlighted the relevant changes in both of these documents. Both are attached to this email.

Please let us know if there is anything else we need to do in order to apply for this deviation.

Many thanks,
Kate Southwood
on behalf of Veerle Vijverberg and myself

****

Catherine L. Ward, Ph.D.
Professor
Interim Head of Department: Social Development
Room 2.09, Department of Psychology
PD Hahn Building
Chemistry Mall, Upper Campus
Rondebosch 7701
University of Cape Town
South Africa
Office: +27-21-650-3422
Mobile: +27-84-601-2244
E-mail: Catherine.Ward@uct.ac.za / Cathy.Ward.SA@gmail.com

Skype: cathy.ward.sa

http://www.psychology.uct.ac.za

Tue, Sep 7, 10:07 AM

Kate Southwood
<katesouthwood96@gmail.com>

to Catherine

Hi Cathy

Thank you so much! We accept those changes.

Kind regards,
Kate

Appendix I - Informed consent form for SRPP participants

Study Title: Perceptions of crimes survey

Purpose

You are invited to take part in a research study which aims to investigate your perceptions of crimes.

Procedure

- By agreeing to take part in this research study, you will be required to watch a video and then answer some questions.
- This procedure will take approximately 10-15 minutes.
- You may withdraw from the study at any time, without explanation required. If you don't want to answer a particular question, then you would need to withdraw from the study. However, the SRPP point will only be awarded on the completion of this study.

Inclusion criteria

You are eligible to participate in this study if you:

- Are a UCT student
- Are 18 years of age or older
- Have not participated in any study relating to crime in the last 18 months.
- Have access to a laptop or desktop computer and stable internet connection to complete this survey

Benefit

You will receive one SRPP point for your participation in this study.

Risk

The researchers have not identified any risks that may result in participating in this study, however, should you experience any distress or concern throughout your participation you may alert the research team who will put you in touch with Student Wellness.

Confidentiality

All responses will be completely confidential. Your personal information (your name and student number) is only being collected for the sake of SRPP point compensation. All your answers to questions that are recorded during the study will not be matched to your identity. They will be coded with participant numbers (no longer attached to your identity) and those anonymous answers will comprise the data for our study. We will be using the data for a thesis and possibly reports and conference presentations: these will be written in such a way that no-one can identify you from the report or presentation. The data will be stored on an encrypted, password protected file, which only the researchers will have access to.

I have read the above explanation of this research study and I am aware of the criteria for inclusion. My questions and concerns have all been answered to my satisfaction. I acknowledge that my participation in this study is completely voluntary, and that I may withdraw my participation from the study at any time. I understand that the information obtained in this study will remain confidential and internal to the study. I have been assured that any questions or concerns I may have in the future will be answered by the researchers.

Name of Participant Student Number

_____ _____

Course Code Date

_____ _____

By giving consent I state that I understand English (can read and write it) and that I have normal, or corrected-to-normal vision (such as wearing glasses or contact lenses if I need them).

- I consent to participate in this study
- I do not consent to participate in this study

Participants may only move forward with the study if they select the "I consent to participate in this study" option. Should they decline, they will be redirected to the closing page of the study, and they will not be rewarded with an SRPP point.

Appendix J - Informed consent form for externally recruited participants

Study Title: Perceptions of crimes survey

Purpose

You are invited to take part in a research study which aims to investigate your perceptions of crimes.

Procedure

- By agreeing to take part in this research study, you agree to watch a video and then answer some questions.
- This procedure will take approximately 10-15 minutes.
- You may withdraw from the study at any time, without explanation required. If you don't want to answer a particular question, then you would need to withdraw from the study. However, you will only be entered into the raffle (see below) on the completion of this study.

Inclusion criteria

You are eligible to participate in this study if you:

- Are 18 years of age or older
- Have not participated in any study relating to crime in the last 18 months.
- Have access to a laptop or desktop computer and stable internet connection to complete this survey

Incentive

Your name will be entered into a raffle to win one of three prizes. There is one prize valued at R500, and two prizes valued at R250 each. Raffle winners will be randomly selected by an independent party. If you are randomly selected to win one of these prizes, you will be contacted via the provided email address. If you do not reply within one week, you will

forfeit your prize and another name will be selected. Note: entry into this raffle does not guarantee winning a prize.

Risk

The researchers have not identified any risks that may result in participating in this study, however, should you experience any distress or concern throughout your participation you may alert the research team who will put you in touch with Student Wellness.

Confidentiality

All responses will be completely confidential. Your personal information (your name and email) is only being collected for the sake of raffle entry, and this information will be stored separately from your answers to our other questions. All your answers to questions that are recorded during the study will not be matched to your identity. They will be coded with participant numbers (no longer attached to your identity) and those anonymous answers will comprise the data for our study. We will be using the data for a thesis and possibly reports and conference presentations: these will be written in such a way that no-one can identify you from the report or presentation. The data will be stored on an encrypted, password protected file, which only the researchers will have access to.

I have read the above explanation of this research study and I am aware of the criteria for inclusion. My questions and concerns have all been answered to my satisfaction. I acknowledge that my participation in this study is completely voluntary, and that I may withdraw my participation from the study at any time. I understand that the information obtained in this study will remain confidential and internal to the study. I have been assured that any questions or concerns I may have in the future will be answered by the researchers.

Name of Participant

_____

Participant Email Address

_____

I have read the above explanation of this research study and I am aware of the criteria for inclusion. My questions and concerns have all been answered to my satisfaction. I acknowledge that my participation in this study is completely voluntary, and that I may withdraw my participation from the study at any time. I understand that the information obtained in this study will remain confidential and internal to the study. I have been assured that any questions or concerns I may have in the future will be answered by the researchers.

- I consent to participate in this study
- I do not consent to participate in this study

Participants may only move forward with the study if they select the "I consent to participate in this study" option. Should they decline, they will be redirected to the closing page of the study, and they will not be rewarded with an entry into the raffle.

Appendix K - Debriefing form

Study title: Similarity within and across single- and multiple-suspect lineups

Thank you for your participation in the study. The study was about finding the optimal model of lineup for police lineup identification accuracy, as well as the optimal levels of similarity, both between suspects and perpetrators, and between lineup members. The study aims to contribute to a field of growing research to ensure higher rates of accurate identifications, and lower rates of false or non-identifications.

How was this tested?

You were asked to watch a video, complete a few questions and tasks, then make an identification (or non-identification) of the perpetrator from a lineup(s), based on the video you watched at the beginning. You were randomly allocated into a condition with a specific style of lineup (in one case, you would have seen one long lineup (called a multiple-suspect lineup), whereas in another condition you would have seen two smaller lineups (two single-suspect lineups). The lineups also had varying levels of similarity between the perpetrator and the suspects, as well as across the lineup members. Your responses were recorded and scored on their accuracy, to determine which lineup model and level of similarity was optimal for correct identification of a perpetrator. Your answers to the questionnaire at the end were also recorded, to investigate which lineup model and level of accuracy might elicit the most confidence, or quickest identification.

The reason you were deceived about the nature of this study (i.e., not told about the lineup from the beginning) was to ensure proper encoding of the crime. Eyewitnesses do not know they are about to witness a crime, so they are not thinking about making a lineup identification when they see the crime take place. We needed to ensure that was the case for the participants too. If you had known a lineup was going to take place, you may have paid more attention to the perpetrator’s face than you otherwise would have. Please note the video of the crime was staged; no real crime was committed.

Hypothesis

Based on existing research, the multiple-suspect model (one longer lineup with several suspects in it) should yield higher accuracy ratings than the single-suspect lineup model (two shorter lineups). Regarding the level of similarity, the lower levels of similarity should yield higher accuracy. Thus, it follows that the multiple-suspect model with low similarity levels should yield the highest accuracy and lowest inaccuracy scores.

Importance of this topic of study

Lineup identification accuracy is vitally important, and thus, finding the model type and similarity levels which optimise accuracy is also important. This would be the model that minimizes the harm caused by false identifications (innocent suspects being criminalised) or lack of identification (failure to identify the perpetrator, leading to perpetrators going free). There is very little research on the multiple-suspect lineup model, although it presents as a promising alternative to single-suspect- and all-suspect lineup models. As there are more suspects present than in a single-suspect lineup, there is a higher chance that a potentially guilty suspect will be identified. Through the inclusion of foils (lineup members who the administrator knows are “innocent”), if one of them is falsely identified, they will not be prosecuted. This model thus appears to maximize the benefit of both single- and all-suspect models, while reducing the associated risk of each.

Further, similarity between lineup members is an area of research that ought to be explored. Although this is a well-informed topic of research, this requirement has not been investigated in relation to the type of lineup model before. Additionally, though similarity is legally required for conducting a multiple-suspect lineup in South Africa, this level of similarity remains undefined, or even which level would be optimal. Thus, there exists a gap in the literature regarding which lineup model is most accurate, and whether the level of similarity between lineup members moderates any advantage one model may have over another.

For more Information

If you would like to learn more about the topic, feel free to consult the following article, which outlines some of the key advantages and disadvantages of the various lineup model types:

Wells, G. L., & Turtle, J. W. (1986). Eyewitness identification: The importance of lineup models. *Psychological Bulletin, 99*(3), 320-329. https://doi.org/10.1037/0033-2909.99.3.320

If you are interested in a report of the findings of this study once it is complete, or if you have any additional questions, please contact Kate Southwood (STHKAT006@myuct.ac.za) or Veerle Vijverberg (VJVVEE001@myuct.ac.za).

If you are experiencing any distress or concern regarding the subject matter of the study, please do not hesitate to make an appointment with Student Wellness. This can be done through the contact details below.

UCT Student Wellness

Website link

Phone: 021 650 1017/20

We hope that you enjoyed contributing to the study and thank you again for your participation.

Appendix L - Result of Binary Logistic Regression Output

Table 3

Logistic Regression Model: Binary Outcome Variable

Effect	Estimate	*SE*	95% CI		*p*
			LL	*UL*	
Target Presence	-2.11	.60	-3.35	-.98	<.0001**
Suspect-Perpetrator Similarity	−.21	.58	−1.37	0.92	.71
Lineup Member Similarity	.42	.62	−0.78	1.67	.49
Lineup Type	−1.25	.56	−2.4	-.17	.03 *
Target Presence* Suspect-Perpetrator Similarity	.62	.83	-1.01	2.26	.46
Target Presence * Lineup Member Similarity	.52	.84	-1.14	2.18	.54
Suspect-Perpetrator Similarity * Lineup Member Similarity	.02	.86	-1.67	1.7	.98
Target Presence * Lineup Type	0.52	0.89	-1.26	2.27	.56
Suspect-Perpetrator Similarity * Lineup Type	1.42	.81	-.15	3.04	.08
Lineup Member Similarity * Lineup Type	.88	.84	-.77	2.53	.29

Target Presence * Suspect-Perpetrator Similarity * Lineup Member Similarity	-.48	1.17	-2.77	1.81	.68
Target Presence * Suspect-Perpetrator Similarity * Lineup Type	-1.25	1.23	-3.66	1.18	.31
Target Presence * Lineup Member Similarity * Lineup Type	-1.09	1.23	-3.49	1.33	.37
Suspect-Perpetrator Similarity * Lineup Member Similarity * Lineup Type	-.85	1.20	-3.2	1.5	.48
Target Presence *Suspect-Perpetrator Similarity * Lineup Member Similarity * Lineup Type	1.24	1.70	-2.11	4.58	.47

Note. df always = 1. *$p < .05$, **$p < .01$, ***$p < .001$. CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

Appendix M - Coefficients of Multinomial Logistic Regression Output

Table 4

Logistic Regression Model: Multinomial Outcome Variable

Effect	Identification Type			
	Correct Rejection	False Identification	False Rejection	Foil Identification
Lineup Type	-.07 [-1.08, .93]	.38 [-1.14, 1.91]	.74 [0.13, 1.34]	0.28 [-.55, 1.11]
Suspect-Perpetrator Similarity	-.38 [-1.39, 0.63]	-1.64 [-3.4, 0.11]	-.2 [-.8, .41]	-.62 [-1.46, .21]
Lineup Member Similarity	-.19 [-1.22, .85]	1.62 [-.6, 3.84]	-.64* [-1.24, -.03]	-1.07* [-1.92, -.22]
Target Presence	-31.55 [-431.32, 368.23]	-16.48 [-187.69, 154.74]	.57 [-397.3, 398.44]	-16.31 [-187.52, 154.9]
Constant	16.94 [-154.27, 188.15]	12.51 [-158.71, 183.73]	0.13 [-397.75, 397.99]	16.30 [-154.91, 187.51]
Akaike Inf. Crit.	804.68	804.68	804.68	804.68

Note. df always = 1. *$p < .05$, **$p < .01$, ***$p < .001$. Values indicate coefficients.

Confidence intervals indicated by square brackets; [lower limit, upper limit].

Appendix N - Factor Loadings of Metacognition data

Metacognition Questionnaire Items.

	Factor 1 Elimination Process	Factor 2 Automatic Recognition	Factor 3 Familiarity-search	Factor 4 Sense of Difficulty
I immediately recognised the perpetrator, but I cannot explain why.	**.76**	-.13		-.18
When I viewed the lineup, I immediately recognised the perpetrator because their face jumped out at me.	**.81**	-.26		-.16
I compared the photographs with each other to make my choice.		**.54**	.23.	.2
I used a process of elimination until I arrived at a final choice.	-.15	**.67**	.16	
I had no sense of familiarity when I looked at the lineup, but I know I would have recognised the perpetrator if they had been presented.	-.15	.22	**.35**	**.59**
I was afraid of not making the right decision.	-.24	.21	.11	**.62**
I tried to remind myself of the perpetrator so that I could compare him to the proposed images.		.13	**.85**	.11
I tried to match the image of the perpetrator that I had in my mind with the images presented.			**.88**	
I would have had less difficulty in making my choice if the people had been presented to me in person rather than through images.		.1	.18	**.33**
I immediately recognised someone, then I compared the other images to be sure.	**.80**	.15		
I first spotted the image that seemed right to me then I looked at the others and gradually eliminated them.	**-.42**	**.62**		.16
I had made a first choice, but I eventually chose another person.	**-.32**	**.58**		.12
I imagined each member of the lineup in the theft situation to determine if they were the perpetrator.		**.58**		
During the task, I was aware that my memory is not infallible and that I could be mistaken.	-.19			**.75**
Some of the images were too similar and it made me doubtful.	**-.36**	**.36**		**.46**
Proportion Variance explained	.16	.31	.12	.12

Note. Coefficients Smaller Than |0.1| are not reported for clarity