The Plymouth Student Scientist - Volume 15 - 2022 The Plymouth Student Scientist - Volume 15, No.2 - 2022

2022

The effect of pre-event instructions on eyewitness identification

Allen, J.

Allen, J. (2022) 'The effect of pre-event instructions on eyewitness identification', The Plymouth Student Scientist, 15(2), pp. 506-523. https://doi.org/10.24382/82k9-0x76
http://hdl.handle.net/10026.1/20126

https://doi.org/10.24382/82k9-0x76
University of Plymouth

The effect of pre-event instructions on eyewitness identification

Jessica Allen

Project Advisor: Dr Mila Mileva, School of Psychology, University of Plymouth, Drake Circus, Plymouth, PL4 8AA

Abstract

Eyewitness identification research has informed policies and procedures, but experiments on eyewitness identification often use unrealistic testing methods. Part of these unrealistic methods are pre-event instructions given before the participant is exposed to the crime. Instructions can direct attention, thus lead to better accuracy. If previous research had used pre-event instructions, results may have overestimated people's ability to identify culprits correctly and called into question the validity of previous conclusions. This would not be an issue if usage of pre-event instructions were reported in the literature, as comparisons could be made between studies that did use them or between studies that did not. Due to the inconsistent reporting of pre-event instructions, the true impact of pre-event instructions on eyewitness accuracy and confidence was previously unknown, a gap the current study aimed to fill. A between-subjects experiment was conducted. 623 participants were randomly assigned either detailed or vague pre-event instructions, viewed a crime video, and identified the culprit from a lineup and gave their confidence rating. Receiver operating characteristic analysis revealed no significant difference in accuracy levels between both conditions. Confidence-accuracy characteristic analysis revealed detailed pre-event instructions resulted in more trustworthy suspect-identifications. The implications of these findings for both the field of eyewitness identification and the justice system are discussed, with suggestions for further research.

Keywords: Eyewitness identification, pre-event instructions, lineup task, attention, ROC analysis, CAC analysis, accuracy, confidence.

Introduction

Eyewitness testimony (EWT) is highly valuable to police officers (Kebbell & Milne, 1998) as it can be one of the first avenues used to establish the events of a crime and the identity of the perpetrator. EWT also heavily influences jurors in their decisions of guilt (Penrod & Cutler, 1995), and incorrect identifications by eyewitnesses can lead to the incarceration of innocent people (Wells, 1993). An eyewitness' confidence in their identification can also impact the persuasiveness of their testimony, regardless of the accuracy (Wells et al., 1979). Therefore, it is important to understand the factors that influence an eyewitness' ability to accurately and confidently identify a suspect. By knowing what impacts identification in controlled settings, it allows for measures to be taken to counteract issues in the real world. As such, research into EWT has informed procedures and policies that shape improvements in the validity and reliability of EWT within the criminal justice system.

Psychology has for years explored and attempted to improve reliability issues regarding EWT. Although the topic has been researched as far back as the early 19th century (Wells et al., 2000), it was the 1970s onward when the majority of research began (Leach et al., 2009). However, it took until the 1990s for the legal system to appreciate and implement the findings from EWT research (Wells, 1993). Research into numerous elements of EWT, such as how people remember faces (Davies et al., 1981), have furthered understanding of the best ways to approach these issues both in research and the real world. Other research areas sparked debate: the validity of repressed memories (e.g., Loftus & Ketcham, 1994; Scoboria et al., 2016), the misinformation effect (e.g., Loftus & Loftus, 1980; Blank & Launay, 2014), and sequential versus simultaneous lineup practices (e.g., Lindsay & Wells, 1985; Mickes et al., 2012), all creating the need for further research to improve understanding.

However, despite the expansion of EWT research, one area that has been neglected is the circumstances that capture attention prior to a crime (Jaeger et al., 2017). People overestimate their ability to notice unusual events as they believe their attention will be caught when the event takes place (Levin et al., 2000). However, criminal activity rarely engages attention as people are often engaged in their own activities and not seeking out a crime to notice it happening (Hyman et al., 2018). Aside from roles with specific training regarding orienting attention to a crime or roles that receive advance warning of a crime occurring, such as police officers or bank workers, people are not usually aware that a crime is happening or where it is important to direct their attention. Nevertheless, experiments using participant-witnesses tend to show either a pre-recorded video or a live performance of a staged or simulated crime that is often in frame and in focus, lacking ecological validity. Taken on its own, this is not egregious. This methodology allows for manipulations that can establish interpretable results (Malpass & Devine, 1980), and produces false identifications at very high rates that echo the real-world phenomenon of witnesses incorrectly identifying an innocent suspect (Wells, 1993).

One element of this experimental methodology, however, is missing. Despite results following similar real-world patterns of behaviours, the staged crimes and laboratory settings cannot replicate real situations (Malpass & Devine, 1980). When compounded with experimenters providing pre-event instructions to participants, such as how to attend to the event or what will be expected of them after the event, the ecological validity is even worse. Further, the inconsistent use of pre-event instructions is itself problematic. When pre-event instructions are used by some experiments and not others, when those that do use them have varying degrees of detail to them, and when experimenters rarely report whether pre-event instructions were used and to what extent, the ecological validity falls apart, regardless of whether results mirror reality. Using unrealistic replications of crimes and providing instructions about the event and/or task can result in participants greatly attending to the event and increase their identification accuracy (Leach et al., 2009).

This inconsistency in reporting pre-event instructions (Baldassari et al., 2021) would not be an issue if it was known exactly how these pre-event instructions impact participant performance. However, no study has purely tested pre-event instructions as a possible moderator for identification accuracy or confidence within the field of EWT. By not knowing whether pre-event instructions impact performance, it calls into question the validity of previous results and conclusions about the accuracy of eyewitness identification. If pre-event instructions do impact eyewitness accuracy and confidence, previous studies that used them may have overestimated identification levels. Even with studies that have not used pre-event instructions, their results cannot be compared to studies that have used them as pre-event instructions may be a confounding variable. Hence the need for the current study, which will be a pure test of the effect of pre-event instructions on eyewitness identification.

Despite little research specifically into pre-event instructions, some research has been done into the impact of other instructions on eyewitness accuracy, with mixed results. Cowan et al. (2014) provided half of their participants with a warning mid-way through interacting with the target that they would be tested on their identification accuracy with a forthcoming lineup. The other half received no such warning. Two weeks after the event, participants were presented with both a culprit-absent (CA) and culprit-present (CP) lineup and asked to identify the culprit in each. Using both CA and CP lineups allows for examination of the impact independent variables can have on both correct and false identifications (Leach et al., 2009). Cowan et al. (2014) found that being warned about the lineup improved identification accuracy but had no impact on confidence. This suggests that having some form of instructions can impact an eyewitness' accuracy but not confidence. However, this study does not reflect the procedure real-world eyewitnesses experience. Usually, eyewitness experiments are a between-subject design, but Cowan et al. (2014) implemented a within-subject design by having participants complete both a CA and CP lineup. Real eyewitnesses do not undergo such repeated measures (Wells, 1993; Mansour et al., 2017), and so by choosing to use this methodology, Cowan et al.'s (2014) study lacks some ecological validity regardless of using instructions.

In contrast, an earlier study by Yarmey (2004) did use a between-subjects design. Similar to Cowan et al. (2014), pre-event instructions were not used but given mid-way through interacting with a target. Participants were either told it was important to remember the target's face for a later memory test or not given any such instruction. Unlike other studies, including Cowan et al. (2014), they found that telling a participant to remember the target had no impact on identification accuracy (cf. Maass & Brigham, 1982). Even though identification accuracy was not affected by instructions, Yarmey's (2004) results did show that instructions impacted participant recall of the target; clothing characteristics were better recalled than physical characteristics compared to the control condition. In terms of confidence, a significant negative correlation was found between accuracy and confidence of identifications when collapsed across conditions, unlike Cowan et al. (2014) where no difference was found. Therefore, both Cowan et al. (2014) and Yarmey (2004) show that instructions affect participant performance and confidence, albeit in different ways. However, as neither study used pre-event instructions, the question of their impact on eyewitness identification accuracy and confidence remains.

One study that did use pre-event instructions was Lindsay et al. (1998), however they were focusing on how accuracy-confidence differed across conditions with multiple variables and so did not separate pre-event instructions to provide a pure test of their impact. Participants were divided into four groups; the "worst condition" watched a very short video of the crime which had camera angles that made it difficult to see the culprit, and were told they would need to identify the filming location; the "medium condition" used the same video, but instead told participants to study the culprit's appearance as they would be asked to identify them later; the final two conditions used longer videos with better angles and the same pre-event instructions as the medium condition (Lindsay et al., 1998, p. 216). They found that participants in the worst condition had lower identification accuracy rates and lower confidence levels than the other conditions. However, as pre-event instructions were confounded with camera angles and view time, it is difficult to establish a clear conclusion as to the impact pre-event instructions alone had on identification accuracy and confidence (Baldassari et al., 2021). Therefore, accounting for the previously discussed literature, not only has there not yet been a true test of the impact of pre-event instructions on eyewitness identification accuracy, it is also not clear how they impact confidence levels either. Consequently, there is a clear gap in the literature and need for an experiment to purely test the impact of pre-event instructions on eyewitness identification accuracy and confidence.

One test of pre-event instructions was Wulff and Hyman's (2021) study, but they manipulated pre-event instructions as a means to test the pervasiveness of inattentional blindness. Inattentional blindness is a failure to notice and become aware of objects that pass through someone's visual field (Mack & Rock, 1998). When applied to EWT, inattentional blindness is a failure to notice a crime, despite it happening in one's field of view, because one's focus is on something else in the environment (Hyman et al., 2018). In Wulff and Hyman's (2021) study, participants were shown a video of a busy hallway with several actors entering and exiting the frame, and at one point a man steals a brightly coloured backpack. Participants were split three conditions: the control condition were told to watch the video; the inattentional blindness condition were told to count the number of people wearing white; and the eyewitness condition were told to watch for a theft. All participants were then asked whether they noticed the theft, to describe the culprit, and attempted to identify the culprit from either a CA or CP lineup. Results showed fewer participants noticed the crime in the control condition than in the eyewitness condition, but identification of the culprit was not affected by condition. However, participants in the eyewitness condition were more likely to identify an innocent bystander as the culprit when both culprit and bystander were present. This suggests that pre-event instructions could have a negative impact on eyewitness identification; increasing awareness of an upcoming crime may result in participants focusing on the crime itself, rather than the person committing it. Additionally, watching for a crime led people to identify the innocent bystander; familiarity from seeing the bystander during the crime led to them being identified as the guilty suspect (Read et al., 1990).

However, there are a few points to note with the study and its generalisability to the field of EWT and pre-event instructions. Wulff and Hyman's (2021) video features a crime that is relatively hard to notice among several other actors walking about the scene. Although possibly more accurate to actual criminal events, it cannot necessarily be compared with most other eyewitness identification studies that use much more conspicuous crimes (see Wells et al., 2020 for a review). Another issue is that the pre-event instructions in this study simply told some of the participants to watch for a crime and not about the lineup task. Therefore, it cannot be assumed that performance in the lineup task was directly caused by these instructions as participants were unaware they would need to identify the culprit from a lineup.

As shown by Wulff and Hyman (2021), pre-event instructions can increase attention towards hard-to-notice crimes. However, whether this effect remains during more obvious crimes is yet to be seen. Although inconspicuous crimes are likely more realistic to what is experienced by witnesses in the real world, conspicuous crimes are usually used during research. Therefore, it is important to know whether pre-event instructions impact conspicuous crimes to allow for comparisons with other literature that used pre-event instructions and enables us to assess the validity of those studies when applying them to real-world situations.

This study aims to fill the gap in the literature about the impact of pre-event instructions on eyewitness identification accuracy and confidence. Based on the reviewed literature, this study hypothesises: participants who receive detailed pre-event instructions will correctly identify the culprit more than participants who receive vague instructions; participants in the detailed condition will reject CA linups more often than participants in the vague instructions condition; and participants in the detailed condition will have better calibrated confidence ratings. These hypotheses will be tested by manipulating pre-event instructions alone by providing either vague (i.e., "Watch this video.") or detailed (i.e., "Watch this video of a crime. You will be asked if you can identify the criminal from a lineup later.") instructions to participants before they watch a video of a crime, complete a lineup identification task, and rate their confidence in their decision. Analysis will compare identification accuracy and confidence across both conditions.

Method

Participants

A total of 623 participants took part in the experiment. 85 participants were recruited using the online SONA system from Stage One and Stage Two of the University of Plymouth's psychology degree programme, in exchange for course credit. All participants were over the age of 18, and as the study was online, had normal or corrected-to-normal vision. There were 502 females, 113 males, and 8 other, with a mean age of 23.92 years (*SD* = 10.2, range 18-70). The participants were predominantly Caucasian (79.45%). The remaining participants identified as Hispanic (3.85%), Black (3.69%), Native American/First Nation (3.21%), East Asian (1.77%), Southeast Asian (1.61%), and other races (6.42%).

Design

The study utilised pre-event instructions as a between-subjects variable. Participants were either given vague (i.e., “Watch this video.”) or detailed (i.e., “Watch this video of a crime. You will be asked if you can identify the criminal from a lineup later.”) pre-event instructions. Both versions of instructions oriented attention to the subsequent video, however the detailed instructions informed participants about the crime in the video and the task that followed. Half the participants viewed a CP lineup and the other half a CA. The primary measures were correct identification (ID) rates of the guilty suspect, false ID rates of innocent suspects, and confidence in those identifications.

Materials

Event Video: Participants viewed a six-second-long video of a man walking into an office, taking keys from a desk, entering a car using the stolen key, and starting the engine. The man is the only person shown on screen, and he is the focus for the entire video. Two versions of the same crime video were used with a different culprit in each and the two culprits were description-matched (Caucasian, light brown/blonde hair, medium build, ~20 years of age). The identity of the culprit was counterbalanced across participants so only one video was viewed, and an equal number of responses were collected for the two culprit identities.

Lineups: Photographs used to create the lineups were taken from various face databases. CA and CP simultaneous lineups were created, and the designated innocent suspect was the criminal from the other video. The filler photos description-matched each culprit’s appearance. Photos were presented in colour, cropped at the neck to avoid clothing clues, had the same lighting conditions, and the identities in all photos were looking directly at the camera. The lineup photos were approximately 371 x 383 pixels in size and were presented in a 3x2 format and numbered from 1 to 6 based on their position.

Apparatus: The experiment was built using the online platform Qualtrics, which also collected all the data. Attention checks were used throughout the experiment to ensure participants were paying attention to the online study.

Procedure

As the study was online, participants were asked to pass bot checks before providing demographic information. They were then posed vigilance questions that asked for their favourite food and required them to retype exactly the main experiment instructions. Participants were randomly allocated to one of two different pre-event instructions conditions: vague (“Watch this video.”) or detailed ("Watch this video of a crime. You will be asked if you can identify the criminal from a lineup later."). Participants then watched a randomly assigned video, answered questions about the colour of the man's top and the car he got into, and then moved onto a distractor task consisting of 15 mental rotation trials. Once completed, participants were presented with a lineup of 6 images. Participants were randomly assigned to see either a CP or CA lineup and were told about the possibility of the actual culprit not being present. Following the lineup presentation was a multiple-choice question asking which image, if any, was the culprit, and a scale for participants to rate the confidence of their choice from 1-7. After, participants were asked questions regarding the experiment itself, then presented with a debrief explaining the aim and rationale of the study. The total time to complete the study was under 15 minutes.

Results

All analyses were performed using R Statistical Software (v4.1.0; R Core Team, 2021). Using the pwr package (v1.3-0; Champely, 2020) a post-hoc statistical power analysis was conducted to test the power between two independent group means using a two-tailed test. The detailed instructions condition had n = 313 and the vague instructions had n = 310. Using a medium effect size (d = 0.5) and significance level of .01, the results showed this research achieved a power of 0.99.

Identification Rates

To show the spread of data, raw ID values for each level of confidence for the detailed and vague instructions condition are reported in Tables 1 and 2 respectively. Comparing Tables 1 and 2 shows that participants in the detailed instructions condition identified foil picks as the culprit (total n = 113) more than participants in the vague instructions condition (total n = 70). Relatedly, participants were poorer at correctly rejecting CA lineups in the detailed instructions condition (total n = 37) than participants in the vague instructions condition (total n = 52), which does not support the hypothesis that participants in the detailed pre-event instructions condition would reject CA lineups more often than participants in the vague instructions condition. Overall, performance for correctly identifying the culprit was poor across all confidence levels in both the detailed instructions (total n = 39) and vague instructions (total n = 42) conditions, representing a correct ID rate of less than 14% for each condition.

Table 1: Raw Identification Rates of Detailed Instructions Condition for Each Level of Confidence

Confidence	False IDs	Foil picks	Misses	Correct rejections	Correct IDs
1	0	12	5	3	2
2	5	10	4	8	4
3	3	26	6	9	6
4	2	25	7	7	11
5	1	28	15	7	11
6	0	9	2	2	4
7	1	3	1	1	1

Table 2: Raw Identification Rates of Vague Instructions Condition for Each Level of Confidence

Confidence	False IDs	Foil picks	Misses	Correct rejections	Correct IDs
1	3	7	8	11	1
2	2	11	4	8	7
3	5	19	8	5	14
4	4	19	11	7	10
5	3	19	12	11	6
6	0	3	1	7	4
7	0	2	4	3	0

Decision Times

Although not a focus of this study, another way of assessing the accuracy of identifications is by looking at response times. The faster an identification is made, the more likely it is to be accurate (Dunning & Perretta, 2002; Sporer, 1993; Weber et al., 2004). To examine whether instructions impacted the time it took for participants to submit their identification, the mean of the submission times were calculated for each condition and trial type (Table 3). Participants took slightly longer to respond to CA trials for both conditions. Across both trial types, participants in the detailed instructions condition took slightly longer to make a decision than those in the vague instructions condition, although response times for both conditions were similar. However, the range of response times for each condition greatly varied. Participants in the vague instructions condition seemed to be fairly similar across both CA and CP trials in how quickly they made a decision regarding the lineup. In the detailed condition, there was a far greater range of response times, particularly for the CA trials, meaning the detailed instructions condition could have produced less accurate responses. To test whether this was the case, receiver operating characteristic analysis was used.

Table 3: Raw Identification Rates of Vague Instructions Condition for Each Level of Confidence

Condition	Culprit absent trials		Culprit present trials	
	M	*SD*	*M*	*SD*
Detailed instructions	26.28	54.44	24.82	22.05
Vague instructions	24.66	15.28	23.95	12.89

Receiver Operating Characteristic (ROC) analysis

ROC analysis measures discriminability (Wixted & Mickes, 2012), or a participant's ability to discriminate an innocent suspect from a guilty one. It can inform policymakers about system variables that influence eyewitness memory (Mickes, 2015); specifically, the ways in which the methods used by the justice and legal systems can influence an eyewitness and their identification accuracy. Other measures used to test discriminability use ratio-based measures (see Wixted et al., 2014 for a review) that can be influenced by a witness' willingness to make a response (Wetmore et al., 2015). In other words, because measure A results in more responses than measure B, it is seen as the better measure because, due to the nature of having more responses and inhibiting response bias, more people identified the suspect correctly. However, measure B may actually have performed better because although less people responded, those that did were more accurate more often than people in measure A. ROC analysis provides a way to explore these potential differences as it accounts for response bias.

In lineup ROCs, only suspect identifications (i.e., designated innocent suspects or guilty suspects) are used to create the curves (Wetmore et al., 2015), meaning data such as correct rejections, misses, and foil picks are not used, which is why neither axis reaches one. Each curve is a plot of correct and false identification rates at each confidence level (rating from 1-7, 7 being the highest confidence level). The curve begins with the proportion of correct IDs from the CP lineup (on the y axis) plotted against the proportion of false IDs from the CA lineup (on the x axis), collapsed across all levels of confidence. The next point of the curve plots the proportion of correct IDs against false IDs for confidence levels 2-7. The next plots the proportion of correct IDs and false IDs for confidence levels 3-7, and so on. Ultimately, the final point plots the proportion of correct IDs against false IDs for confidence level 7, the most conservative identification rate. These points are then joined with a line to produce an ROC curve. The closer a curve bows to the upper-left corner, the better the measure can produce participant discriminability; when comparing two ROCs, the curve that bows closer to the upper-left is the better measure as it produces a better balance between culprit IDs and innocent suspect IDs (Smith et al., 2020).

To test whether the detailed pre-event instructions produced higher correct ID rates and lower false ID rates than the vague instructions, ROC analysis was conducted and the ROC curves are shown in Figure 1. The diagonal line depicts where false ID equals correct ID rate. Both conditions produced discriminability above chance. Although the curves are similar, the detailed condition curve is closer to the upper-left corner; it yielded higher correct ID rates and lower false ID rates. Therefore, the detailed instructions appeared to be the better procedure for discriminability between the guilty culprit and innocent suspect, supporting the hypotheses that detailed pre-event instructions would produce better correct IDs and lower false IDs than vague instructions.

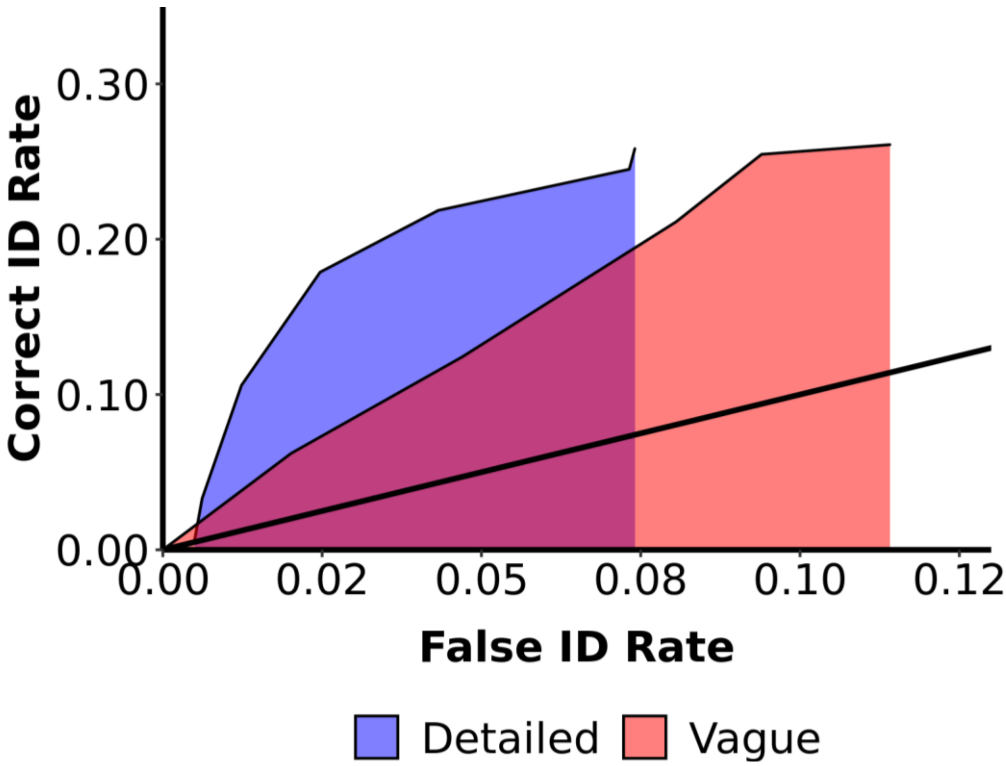

Figure 1: Confidence-Based Receiver Operating Characteristics
Note. The solid black line represents chance performance.

To explore the statistical significance of the two different instructions conditions, partial area under the curve (pAUC) was utilised due to the correct ID and false ID data not forming a full ROC across 0-1. The greater the area under the curve of the ROC, the better participants were able to distinguish innocent and guilty suspects. Measuring the pAUC from the parameters of the shorter, detailed condition curve instead of the larger, vague condition curve meant avoiding assuming and extrapolating data from the shorter curve (Smith et al., 2020), which would have resulted in a less accurate measure. Comparison for both conditions therefore was over the common pAUC region of .99-.89. Partial AUC values were generated using the pROC statistical package (Robin et al., 2011), and statistical analysis was based on the nonparametric bootstrap method (n= 2,000). Results showed that there was not a significant difference in pAUC between the detailed instructions condition (0.014) and the vague instructions condition (0.009), $D = 1.44$, $p = .15$). D is the outcome of (AUC1-AUC2)/s, where s is the standard error of difference between the two pAUCs estimated using the bootstrap method of 10,000 bootstraps (Wetmore et al., 2015). Thus, both detailed and vague instructions produced similar levels of discriminability, which does not support the hypothesis that detailed pre-event instructions caused better correct ID rates and lower false ID rates than vague instructions. However, although both conditions were similar in discriminability, there could still be a difference in trustworthiness across confidence levels, which can only be addressed by confidence-accuracy characteristic (CAC) analysis.

Confidence-Accuracy Characteristic analysis
CAC analysis measures the relationship between the participants' subjective assumption their ID is correct (using a confidence scale from 1-7) and the objective accuracy of their ID. For each level of confidence, suspect ID accuracy was

computed by # correct IDs/(# correct IDs + # incorrect IDs). This makes it an informative measure for jurors and judges, who are interested in whether high-confidence identifications equal high-accuracy (Mickes, 2015). To test whether detailed pre-event instructions gave participants higher confidence when choosing correctly, CAC analysis was conducted. Figure 2 plots identification accuracy for low, medium, and high confidence. This provides insight into the trustworthiness of a participant's judgement, or whether a lineup ID made with high confidence is likely to be accurate. This also gives insight into whether the type of instructions given to participants affected the trustworthiness of their judgements.

The graph demonstrates how the detailed instructions provided more accurate results across all confidence levels compared to the vague instructions. The fact that the accuracy of high-confidence answers in the vague instructions condition do not match the accuracy of high-confidence answers in the detailed instructions condition is unusual. In other studies that used CAC analysis (e.g., Mickes, 2015), participants adjust their confidence depending on the amount of information they have and ultimately both conditions tend to be equally accurate in high confidence judgements. Although the pAUC analysis showed no significant difference in discriminability performance, CAC analysis showed participants in the detailed instructions condition were more likely to make accurate IDs compared to participants in the vague instructions condition. This was especially the case for high confidence ratings; detailed instructions gave more trustworthy IDs than vague instructions, supporting the hypothesis that detailed instructions gave participants higher confidence for correct identifications compared to vague instructions.

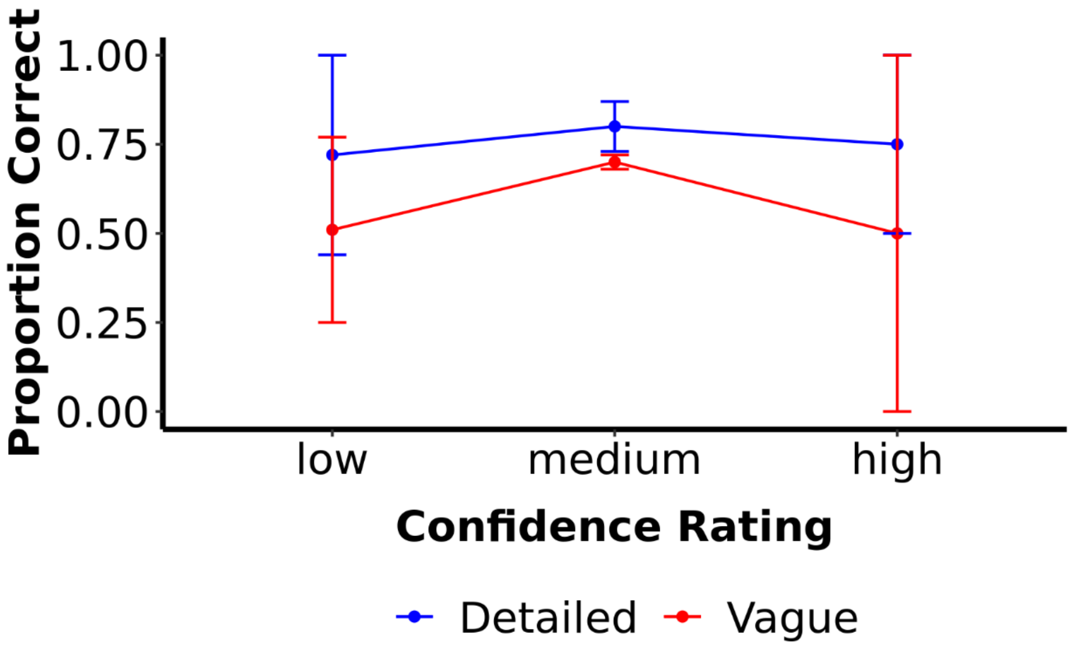

Figure 2: Confidence-Accuracy Characteristic Curves for Each Instructions Condition
Note. Confidence level ratings were aggregated such that ratings of 1-2 were classed as low, ratings 3-5 were classed as medium, and ratings 6-7 were classed as high. Error bars show standard deviation.

Discussion

The current study aimed to provide a novel pure test of the impact of pre-event instructions on eyewitness accuracy and confidence. To conduct this test, detailed and vague instructions were compared. Support was found for the hypothesis about confidence, but not for the hypotheses about accuracy and rejection rates. Unlike what was hypothesised, pre-event instructions did not produce better correct IDs to false IDs. The ROC curves, which calculated discriminability, seemed to suggest detailed instructions produced better identifications rates than vague instructions. However, pAUC analysis showed that this difference in correct ID to false ID rates between the two conditions was not significant, thus suggesting pre-event instructions have no impact on eyewitness accuracy. This validates previous eyewitness identification literature because the use of pre-event instructions is relatively inconsequential to the outcome of the research.

Again, unlike what was hypothesised, detailed pre-event instructions did not improve rejection rates. The raw data showed that detailed instructions led to lower rejection rates of CA lineups and higher levels of foil picks compared to vague instructions, despite both conditions receiving instructions before the lineup task that clearly stated the culprit may not be present. This suggests that detailed pre-event instructions made participants less likely to reject a lineup and more likely to choose someone as being guilty, even when the culprit was absent. These results are reflected in the decision times. Detailed instructions produced a far greater range of decision times: in CP trials, participants differed by over 20 seconds; in CA trials the difference was almost a minute. Both decision times and the raw data spread indicates that detailed pre-event instructions caused participants to take longer to decide the culprit, leading to more inaccurate identifications (Sporer, 1993) and decreased rejection rates. These findings support Wulff and Hyman's (2021) study that found pre-event instructions had a negative impact on identification accuracy.

The rejection of the two hypotheses on accuracy and rejection rates could have several explanations. One is that the participants given detailed instructions did not properly process them. Although all participants had to exactly type out their given instructions, participants may not have completely processed the detailed instructions and instead understood the more general message (i.e., the video will be of a crime, identify the culprit) (Lampinen et al., 2020). This could explain why participants in the detailed condition did not reject the CA lineup as much; they may have assumed they would only be shown a CP lineup. Relatedly, the instruction type given may have affected response bias. Participants given detailed instructions may have felt more willing to claim a culprit was present regardless of whether they were or not (e.g., Stanislaw & Todorov, 1999). Receiving more information could have caused them to feel they had to identify someone, hence the increased foil picks compared to the vague condition. The explanations provided are not necessarily independent from one another. Therefore, even though detailed pre-event instructions may not impact accuracy, they can affect the forms of incorrect identifications a participant makes (i.e., foil picks or rejections), overall making participants more likely to identify someone in a lineup even if the culprit is not present. Although ROC analysis suggests there is no impact on discriminability for pre-event instructions, detailed pre-event instructions seem to impact the ways in which a participant could be incorrect. Therefore, although the current study's results do not invalidate previous conclusions into the accuracy of eyewitnesses, they do suggest that there may be an effect on misidentifications; if previous research used pre-event instructions, then conclusions drawn about the rates of misidentifications could be incorrect.

For confidence, as was hypothesised, participants in the detailed instructions condition had better calibrated confidence ratings. CAC analysis showed the detailed instructions gave more trustworthy IDs than the vague instructions. That is, even though participants' discriminability may be similar for both conditions, participants in the detailed instructions condition were more likely to make accurate IDs compared to participants in the vague instructions condition, especially for high confidence ratings. This finding highlights the importance of conducting both ROC and CAC analysis. Although ROC analysis showed there was no significant difference in discriminability for pre-event instructions, there was a difference in terms of trustworthiness; detailed instructions created more trustworthy identifications. This difference in trustworthiness is a notable outlier compared to other research, which has found that accuracy of high confidence ratings did not differ between conditions (Palmer et al., 2013). Not following this trend suggests that detailed pre-event instructions caused participants to recognise their poor memory and appropriately adjust their confidence, but vague instructions did not. Therefore, as overall accuracy is the same for both conditions, confidence is the most important aspect of an eyewitness' identification. Trustworthiness is important to real-world justice processes as judges and jurors rely on an eyewitness' confidence in their identification to inform their verdict (Penrod & Cutler, 1995; Wells et al., 1979).

The current study was a strong, pure test of the impact of pre-event instructions on eyewitness accuracy and confidence, a focus not previously researched. It used both ROC and CAC analysis to assess the impact of pre-event instructions on both accuracy and confidence, information that is valuable for both policymakers and judges and jurors (Mickes, 2015). It also reinforces the need for researchers to include as much information as possible about the methods used in experiments. If prior research reported their usage of pre-event instructions, perhaps there would be a better understanding of their impact. Although pre-event instructions were not found to impact accuracy, they do impact confidence, setting a precedent for future research to consider the impact of other seemingly inconsequential influences on eyewitness accuracy and confidence, for example the type of crime used in videos or the exact length of time between viewing the crime and identifying the culprit.

The current study is not without its limitations. Firstly, there are issues with ROC analysis. When determining guilt in the real world, correct and false IDs by eyewitnesses are not the only aspects taken into consideration; foil picks and rejection rates also inform guilt, as well as the associated confidence with each decision (Smith et al., 2020). ROC analysis discounts these other considerations and only accounts for identifications. Although the author did account for foil picks and rejections in their own analysis, this was done using raw data so no significance can be drawn. Another problem with ROC is that it discards a lot of suspect IDs (Smith et al., 2020). By using the parameters of the shorter, detailed condition curve, some of the suspect IDs from the vague condition were discarded. Although this allowed for analysis to be conducted on known data, it discounted the higher suspect ID rate that the vague condition had, leading the pAUC analysis to state there was no difference between the two conditions. Using the shorter parameters means it is not known whether those discarded suspect IDs from the vague instructions were correct, so there may be an advantage to using vague instructions.

Another limitation was the length of the crime video. The current video was six seconds long and played with little warning, meaning participants' attention may not have been oriented to the video until it was too late. Although perhaps truer to the experience of real-world crimes, it could have contributed to the low levels of accuracy independent of the pre-event instructions. Further research could be done with different crime video lengths; more exposure to the culprit tends to lead to better accuracy (Leach, 2009), so using different video lengths would assess whether pre-event instructions impact a situation known to lead to better identification accuracy.

Despite not being the focus of the research, the most important finding of the current study was the low rates of correct identifications. Participants were very poor at identifying the guilty culprit, which is concerning given people can be convicted solely on eyewitness identification alone, despite recommendations otherwise (Devlin, 1976). Although the inability of eyewitnesses to correctly identify a culprit is a somewhat accepted phenomenon in the field of eyewitness testimony (Wells, 1993), future research should highlight just how inaccurate people can be. Justice systems need to be aware of the full extent of people's difficulties with correct identifications as innocent people are at risk of being convicted. Relying on something people are unable to do accurately indicates the need for a change in the weight eyewitness identifications hold, regardless of the confidence of such identifications.

Conclusions

The current study's results were mixed, with no clear answer about the impact of pre-event instructions on eyewitness accuracy and confidence. They were not found to impact discriminability but did impact trustworthiness, an interaction future research could explore further. It also validated previous research into eyewitness identification, as research seemingly did not overestimate people's identification accuracy given pre-event instructions relatively inconsequential impact. This study has created avenues for future research to consider, and highlighted people's inability to correctly identify a guilty culprit, a fundamental issue with criminal justice systems' reliance on eyewitness identification.

Acknowledgements

The current project was part of a collaboration with many researchers from several institutions. Special thank you to the following people for realising need for the current research area, designing the experiment, and welcoming data contributions: Kara Moore (Oklahoma State University), Mario Baldassari (University of Saint Francis), Ira Hyman (Western Washington University), Lorraine Hope (University of Portsmouth), Eric Mah (University of Victoria), and Stephen Lindsay (University of Victoria). Special thank you also to the following people for their data contributions; Hannah Rath (Oklahoma State University); Josh Pedretti (University of Saint Francis); Renan Benigno (University of Portsmouth); Ruth Horry, Lauren Kelly, Rosie Jones, Shannan Vale & Bethany Lawson (Swansea University); Andrew Clark (University of Bedfordshire); Jess Karanian (Fairfield University); Hayley Cullen & Zoe Crittenden (University of Newcastle, Australia); Celine van Golde (University of Sydney); Tomás Palma, Francisco Cruz & Joana Quarenta (University of Lisbon); Brittany Jeye (Worcester State University); Ine Van der Cruyssen & Sera Wiechert

(University of Amsterdam); Jamal Mansour (Queen Margaret University); Natalie Mestry (Bournemouth University); Rebecca Helm (University of Exeter); Feni Kontogianni & Kieran Fletcher (University of Winchester); Brian Cahill (University of Florida).

Thank you to my wonderful supervisor Dr Mila Mileva, without whom I would not have had the chance to participate in this research project. If it was not for her incredible generosity with time, knowledge, and encouragement from the very beginning, this project would not have got the point it has. I sincerely appreciate all the help and advice.

To my parents and partner, thank you for everything you do for me. I would not be here without your love and support. Your encouragement to live for myself and belief that I am capable of anything I want have given me the ability to achive.

References

Baldassari, M. J., Moore, K., Hyman, I., Hope, L., Mah, E. Y., & Lindsay, D. S. (2021). *The effect of pre-event instructions on eyewitness identification.* OSF Preprints. https://osf.io/sk5bc

Blank, H., & Launay, C. (2014). How to protect eyewitness memory against the misinformation effect: A meta-analysis of post-warning studies. *Journal of Applied Research in Memory and Cognition*, *3*(2), 77–88. https://doi.org/10.1016/j.jarmac.2014.03.005

Champely, S. (2020). *pwr: Basic Functions for Power Analysis*. R-Packages. https://CRAN.R-project.org/package=pwr

Cowan, S., Read, J. D., & Lindsay, D. S. (2014). Predicting and postdicting eyewitness accuracy and confidence. *Journal of Applied Research in Memory and Cognition*, *3*(1), 21–30. https://doi.org/10.1016/j.jarmac.2014.01.002

Davies, G., Ellis, H., & Shepherd, J. (1981). *Perceiving and remembering faces*. Academic Press.

Devlin, P. (1976). Report to the Secretary of State for the Home Department of the Departmental Committee on Evidence of Identification in Criminal Cases. HMSO. https://www.gov.uk/government/publications/report-to-the-secretary-of-state-for-the-home-department-of-the-departmental-committee-on-evidence-of-identification-in-criminal-cases-1976

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

Hyman, I. E., Wulff, A. N., & Thomas, A. K. (2018). Crime Blindness: How Selective Attention and Inattentional Blindness Can Disrupt Eyewitness Awareness and Memory. *Policy Insights from the Behavioral and Brain Sciences*, *5*(2), 202–208. https://doi.org/10.1177/2372732218786749

Jaeger, C. B., Levin, D. T., & Porter, E. (2017). Justice is (change) blind: Applying research on visual metacognition in legal settings. *Psychology, Public Policy, and Law*, *23*(2), 259–279. https://doi.org/10.1037/law0000107

Kebbell, M. R., & Milne, R. (1998). Police Officers' Perceptions of Eyewitness Performance in Forensic Investigations. *The Journal of Social Psychology*, *138*(3), 323–330. https://doi.org/10.1080/00224549809600384

Lampinen, J. M., Race, B., Wolf, A. P., Phillips, P., Moriarty, N., & Smith, A. M. (2020). Comparing detailed and less detailed pre-lineup instructions. *Applied Cognitive Psychology*, *34*(2), 409–424. https://doi.org/10.1002/acp.3627

Leach, A.-M., Cutler, B. L., & Van Wallendael, L. (2009). Lineups and Eyewitness Identification. *Annual Review of Law and Social Science*, *5*(1), 157–178. https://doi.org/10.1146/annurev.lawsocsci.093008.131529

Levin, D. T., Momen, N., Drivdahl, S. B., & Simons, D. J. (2000). Change Blindness Blindness: The Metacognitive Error of Overestimating Change-detection Ability. *Visual Cognition*, *7*(1-3), 397–412. https://doi.org/10.1080/135062800394865

Lindsay, D. S., Read, J. D., & Sharma, K. (1998). Accuracy and Confidence in Person Identification: The Relationship Is Strong When Witnessing Conditions Vary Widely. *Psychological Science*, *9*(3), 215–218. https://doi.org/10.1111/1467-9280.00041

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

Loftus, E. F., & Ketcham, K. (1994). *The myth of repressed memory : false memories and allegations of sexual abuse*. St. Martin's Press.

Loftus, E. F., & Loftus, G. R. (1980). On the permanence of stored information in the human brain. *American Psychologist*, *35*(5), 409–420. https://doi.org/10.1037/0003-066x.35.5.409

Maass, A., & Brigham, J. C. (1982). Eyewitness Identifications. *Personality and Social Psychology Bulletin*, *8*(1), 54–59. https://doi.org/10.1177/014616728281009

Mack, A., & Rock, I. (1998). *Inattentional blindness*. MIT Press.

Malpass, R. S., & Devine, P. G. (1980). Realism and eyewitness identification research. *Law and Human Behavior*, *4*(4), 347–358. https://doi.org/10.1007/bf01040626

Mansour, J. K., Beaudry, J. L., & Lindsay, R. C. L. (2017). Are multiple-trial experiments appropriate for eyewitness identification studies? Accuracy, choosing, and confidence across trials. *Behavior Research Methods*, *49*(6), 2235–2254. https://doi.org/10.3758/s13428-017-0855-0

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

Mickes, L., Flowe, H. D., & Wixted, J. T. (2012). Receiver operating characteristic analysis of eyewitness memory: Comparing the diagnostic accuracy of simultaneous versus sequential lineups. *Journal of Experimental Psychology: Applied*, *18*(4), 361–376. https://doi.org/10.1037/a0030609

Palmer, M. A., Brewer, N., Weber, N., & Nagesh, A. (2013). The confidence-accuracy relationship for eyewitness identification decisions: Effects of exposure duration, retention interval, and divided attention. *Journal of Experimental Psychology: Applied*, *19*(1), 55–71. https://doi.org/10.1037/a0031602

Penrod, S., & Cutler, B. (1995). Witness confidence and witness accuracy: Assessing their forensic relation. *Psychology, Public Policy, and Law*, *1*(4), 817–845. https://doi.org/10.1037/1076-8971.1.4.817

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

Read, J. D., Tollestrup, P., Hammersley, R., McFadzen, E., & Christensen, A. (1990). The unconscious transference effect: Are innocent bystanders ever misidentified? *Applied Cognitive Psychology*, *4*(1), 3–31. https://doi.org/10.1002/acp.2350040103

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

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*(1). https://doi.org/10.1186/1471-2105-12-77

Scoboria, A., Wade, K. A., Lindsay, D. S., Azad, T., Strange, D., Ost, J., & Hyman, Ira E. (2016). A mega-analysis of memory reports from eight peer-reviewed false memory implantation studies. *Memory*, *25*(2), 146–163. https://doi.org/10.1080/09658211.2016.1260747

Smith, A. M., Yang, Y., & Wells, G. L. (2020). Distinguishing Between Investigator Discriminability and Eyewitness Discriminability: A Method for Creating Full Receiver Operating Characteristic Curves of Lineup Identification Performance. *Perspectives on Psychological Science*, *15*(3), 589–607. https://doi.org/10.1177/1745691620902426

Sporer, S. L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology*, *78*(1), 22–33. https://doi.org/10.1037/0021-9010.78.1.22

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

Weber, N., Brewer, N., Wells, G. L., Semmler, C., & Keast, A. (2004). Eyewitness Identification Accuracy and Response Latency: The Unruly 10-12-Second Rule. *Journal of Experimental Psychology: Applied*, *10*(3), 139–147. https://doi.org/10.1037/1076-898x.10.3.139

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

Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist*, *48*(5), 553–571. https://doi.org/10.1037/0003-066x.48.5.553

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 evidence. *Law and Human Behavior*, *44*(1), 3–36. https://doi.org/10.1037/lhb0000359

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

Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., & Fulero, S. M. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist*, *55*(6), 581–598. https://doi.org/10.1037/0003-066x.55.6.581

Wetmore, S. A., Neuschatz, J. S., Gronlund, S. D., Wooten, A., Goodsell, C. A., & Carlson, C. A. (2015). Effect of retention interval on showup and lineup performance. *Journal of Applied Research in Memory and Cognition*, *4*(1), 8–14. https://doi.org/10.1016/j.jarmac.2014.07.003

Wixted, J. T., & Mickes, L. (2012). The Field of Eyewitness Memory Should Abandon Probative Value and Embrace Receiver Operating Characteristic Analysis. *Perspectives on Psychological Science*, *7*(3), 275–278. https://doi.org/10.1177/1745691612442906

Wixted, J. T., Gronlund, S. D., & Mickes, L. (2014). Policy Regarding the Sequential Lineup Is Not Informed by Probative Value but Is Informed by Receiver Operating Characteristic Analysis. *Current Directions in Psychological Science*, *23*(1), 17–18. https://doi.org/10.1177/0963721413510934

Wulff, A. N., & Hyman, I. E. (2021). Crime Blindness: The Impact of Inattentional Blindness on Eyewitness Awareness, Memory, and Identification. *Applied Cognitive Psychology*, *36*(1). https://doi.org/10.1002/acp.3906

Yarmey, A. D. (2004). Eyewitness Recall and Photo Identification: a field experiment. *Psychology, Crime & Law*, *10*(1), 53–68. https://doi.org/10.1080/1068316021000058379