

University of Arkansas, Fayetteville

ScholarWorks@UARK

Psychological Science Undergraduate Honors
Theses

Psychological Science

5-2022

The Influence of Facial Exposure on Confidence in Eyewitness Identification

Madelynn Fancher

University of Arkansas, Fayetteville

Follow this and additional works at: <https://scholarworks.uark.edu/psycuht>



Part of the [Cognition and Perception Commons](#), and the [Cognitive Psychology Commons](#)

Citation

Fancher, M. (2022). The Influence of Facial Exposure on Confidence in Eyewitness Identification. *Psychological Science Undergraduate Honors Theses* Retrieved from <https://scholarworks.uark.edu/psycuht/26>

This Thesis is brought to you for free and open access by the Psychological Science at ScholarWorks@UARK. It has been accepted for inclusion in Psychological Science Undergraduate Honors Theses by an authorized administrator of ScholarWorks@UARK. For more information, please contact scholar@uark.edu, uarepos@uark.edu.

The Influence of Facial Exposure Duration on Confidence in Eyewitness Identification

Madelynn Fancher

Department of Psychological Science, University of Arkansas

Honors Thesis

Dr. James Lampinen

May 5, 2022

Abstract:

In crimes in which there is an eyewitness identification, confidence is usually a good predictor of accuracy. However, in some cases, estimator variables might affect the relationship between eyewitness confidence and accuracy. This study analyzes the effect of exposure duration on confidence during an eyewitness identification. According to the pristine conditions hypothesis (Wixted & Wells, 2017), if the system variables are optimal, confidence and accuracy will be strongly related, even if the viewing conditions are suboptimal. Participants in this online study viewed a mock crime in one of two conditions: brief exposure or long exposure. Following viewing the crime, participants completed a distractor task before making an identification in a culprit-absent or culprit-present lineup (randomly assigned). Following the identification, participants indicated their level of confidence in their choice. I hypothesized that highly confident participants would be highly accurate in the long duration condition but not the brief duration condition. However this was not the case. Highly confident witnesses were highly confident, regardless of the length of time they were exposed to the perpetrator's face.

Keywords: eyewitness identification, confidence, estimator variables, pristine conditions hypothesis

The Influence of Facial Exposure Duration on Confidence in Eyewitness Identification

Eyewitness identification is a common method to identify suspects in cases of law. Yet, eyewitness memory is often prone to error (Deffenbacher et al., 2008). Mistaken eyewitness identifications were present in about 70% of DNA exoneration cases (Innocence Project, 2009). In laboratory studies, the guilty suspect is only selected about 50% of the time in culprit-present lineups (Clark et al., 2008). People mistakenly identify innocent fillers in around 50% of culprit-absent lineups. (Clark et al., 2008). In field studies of actual police lineups, 30-40% of all lineup choices are an innocent person (Wells, Steblay & Dysart, 2015).

These odds of misidentification are not ideal and have negative ramifications on society. When eyewitnesses identify the wrong person, it can pose a risk to innocent citizens, often costing them long amounts of time in prison (Sonenshein & Nilon, 2010). When eyewitnesses fail to identify guilty persons, it can create a risk to society at large (Lampinen, Smith, & Toglia, 2021). My thesis examines one variable that can affect the accuracy of choices from lineups, the amount of time the witness has to view the perpetrator's face (Bornstein et al., 2012).

Lineup Construction

A lineup is a group of individuals who are shown to a witness to determine if the witness can identify the perpetrator. In the U.S., most lineups contain six people (Police Executive Research Forum, 2013). Police in other countries often include more individuals in lineups (Fitzgerald, Rubinova, & Juncu, 2021). Sometimes lineups are conducted live, where the six people are in person for the eyewitness to select. More commonly there are six photos of people presented to the witness (i.e., a photo-array). Well conducted lineups contain two types of people: suspects and fillers. A suspect is the person who law enforcement thinks is guilty. Fillers are other people in the lineup, who are known to be innocent, but resemble the description of the

suspect. Picking fillers for a lineup is of crucial importance (Colloff et al., 2021). If fillers do not provide plausible alternatives to the suspect, then the suspect is likely to stand out, increasing the risk to an innocent suspect. For instance, if a witness describes the perpetrator as a blonde woman in her twenties, but the fillers are brunette women in their 40s, it will be obvious to the witness who the suspect is.

To construct a lineup, and ensure that fillers provide plausible alternatives, different tactics may be used (Wells, Rydell, & Seelau, 1993). A lineup may be constructed based off of an initial description of the suspect, and people who match this description are selected to be fillers. Sometimes, a lineup is constructed by looking at the photo of the suspect and selecting the fillers who look similar to the suspect without using the description. Sometimes, a lineup can be constructed using both of these methods

In real world investigations, police do not know for sure whether the suspect is guilty or innocent, they merely have a suspicion. In laboratory studies of eyewitness identification, on the other hand, researchers do know whether the suspect is guilty or innocent (Quigley-MacBride & Wells, 2021). Lineups conducted in the laboratory, in which the suspect is guilty, are called culprit present lineups. Culprit present lineups in the laboratory are meant to mimic the situation in the real world where the police have a suspect, and the suspect really is guilty. Lineups conducted in the laboratory, which contain only innocent people, are called culprit absent lineups. Culprit absent lineups in the laboratory are meant to correspond to a situation in the real world where the police have a suspect, but the suspect is innocent. In some laboratory experiments, one of the people in the culprit absent lineup is chosen by the researcher to serve as innocent suspect (i.e., designated innocent suspect). More commonly, none of the people in the culprit absent lineup is designated as an innocent suspect. Instead, to estimate the rate at which

an innocent suspect would be chosen, the researchers take the overall choosing rate from the culprit absent lineup and divide by the lineup size – typically 6. This is called the estimated innocent suspect identification rate.

System and Estimator Variables

Wells (1978) drew a distinction between two types of variables that can influence eyewitness performance. System variables are variables that are able to be controlled by the criminal justice system. These are things like selecting fillers that are fair (Malpass et al., 2007), ensuring a double-blind procedure (Kovera & Evelo, 2017), providing correct pre-lineup instructions (Lampinen et al., 2020), and avoiding feedback to reinforce selections made by the witness (Douglass & Steblay, 2006). Estimator variables are variables that cannot be controlled by the justice system, and are dependent on the situation, environment, and circumstance. These include factors such as lighting (Nyman et al., 2019), distance (Lampinen et al., 2014), and if a weapon was present (Fawcett et al., 2013).

Recently, some scholars have made the argument that estimator variables and system variables are related in an important way. According to the ‘pristine conditions hypothesis’, if police follow all the correct procedures (system variables), then high confidence suspect identifications will be extremely accurate, even if the viewing conditions (estimator variables) are very poor (Wixted & Wells, 2017). Wixted and Wells summed up their position as follows,

However, after more than 30 years of eyewitness-identification research, our understanding of how to properly conduct a lineup has evolved considerably, and the time seems ripe to ask how eyewitness confidence informs accuracy under more pristine testing conditions (e.g., initial, uncontaminated memory tests using fair lineups, with no lineup administrator influence, and with an immediate confidence statement). Under

those conditions, mock-crime studies and police department field studies have consistently shown that, for adults, (a) confidence and accuracy are strongly related and (b) high-confidence suspect identifications are remarkably accurate (Wixted & Wells, 2017, p. 10).

According to their proposal, when witnessing conditions are poor (e.g., bad lighting, presence of a weapon, high stress), there will be fewer high confidence suspect identifications, but the high confidence suspect identifications that do occur will be highly accurate.

The pristine conditions hypothesis is specifically about how confidence and accuracy are related in those cases where the suspect is selected. The hypothesis excludes from consideration the identification of fillers. If someone is in court testifying, the judge and jury is not interested if a filler is identified, they are interested if a witness identifies a suspect, and how likely is it that the suspect is guilty based on these identifications. The relationship between confidence and accuracy of suspect identifications is often displayed in Confidence Accuracy Characteristic (CAC) curves (Mickes, 2015). A CAC shows confidence on the X-Axis and the accuracy of suspect identifications on the Y-Axis.

In the studies reviewed by Wixted and Wells (2017), highly confident witnesses were typically very accurate when they identified a suspect under pristine conditions (typically 95% accurate or higher). For instance, in one study reviewed by Wixted and Wells, participants viewed a crime in which a weapon was visible, was implied, or was absent altogether (Carlson et al., 2017). In all three conditions, high confidence suspect identifications were correct more than 95% of the time. In another study, they reviewed, participants were presented with lineups either one hour or one week after viewing the crime (Juslin et al., 1996). Regardless of the retention interval, high confidence suspect identifications were correct close to 100% of the time.

Altogether, Wixted and Wells (2017) reviewed 19 different studies in which pristine conditions were used, and in every case, highly confident suspect identifications were highly accurate.

However, in some recent studies, results are inconsistent with the pristine conditions hypotheses. For instance, in one experiment described by Nguyen, Pezdek and Wixted (2017), high confidence cross-race suspect identifications were only correct about 70% of the time. In another study, Giacona, Lampinen, and Anastasi (2021) compared a situation in which witnessing conditions were optimal in multiple ways (e.g., close proximity, long duration, low stress) with a situation in which witnessing conditions were suboptimal in multiple ways (e.g., far distance, brief duration, high stress) and found that high confidence accuracy was significantly lower when viewing conditions were poor. Lockamy et al. (2020) found that when viewing distance was 20 meters, highly confident suspect identifications were correct less than 65% of the time.

Giacona et al. (2021) proposed that when viewing conditions get poor enough, witnesses have difficulty adjusting their response criterion in order to maintain high accuracy, even when they are highly confident. To test this explanation, my goal is to compare a situation where witnessing conditions are likely to be very good with a condition where they are likely to be extremely poor. Some participants were shown a video in which the perpetrator's face was visible for a long time and other participants were shown a video in which the perpetrator's face is visible for a very brief time. In a prior study (Memon et al., 2003), a long exposure duration (45 seconds) was associated with higher accuracy and higher confidence than in a brief exposure duration condition (12 seconds). For my study, the brief exposure condition will be even shorter than in Memon et al (2003).

Method

Participants

Paid participants were recruited from the Prolific online data collection site. Participants were paid \$0.67 for their participation. Initially, 2313 participants began the experiment, but 334 were excluded for various reasons, leaving 1979 participants. Demographic characteristics of the sample are shown in Appendix A. Reasons for excluding participants included self-reported technical problems while watching the video (N = 97), missing one or both of the attention check questions after the video (N = 299), or failing to respond to the lineup at all (N=118). The attention check questions were two multiple choice questions, each with four alternatives, that were given immediately after the video. Participants were to select the color of the car and the object that was stolen from the car. I reasoned that anyone who was paying attention to the video would be able to answer these questions. Some people were excluded for multiple reasons.

Giacona and Lampinen (2021) argued that when the purpose of the study is to examine high confidence accuracy, researchers should over-sample from conditions where the number of high confidence judgements is expected to be lower. Because I assumed that the number of high confidence participants in the brief exposure condition will be very low, I assigned 18 people to the brief exposure condition for every 1 person I assigned to the long exposure condition.

Materials

Crime Video. The mock crime was filmed with an actress and actor, a car, and the actress “stole” a pink purse. The film showed a woman walk up to a car with an open window and take out a purse. Off camera a man verbally confronts her and asks if it is her purse. In the brief duration condition, the woman denies that she is stealing the purse and then abruptly walks off. Her face is visible for approximately 7 seconds. In the long duration condition, there is an

extended verbal exchange between the woman and the man who asks if the purse is hers. In the long duration condition, her face is visible for approximately 45 seconds.

Lineup. The lineups used in the present experiment are shown in Appendix B. To create the lineups used in this experiment, I used the description match approach recommended by Luus and Wells (1991). Seventeen participants completed a description task as part of an extra credit assignment for an advanced psychology class. Each participant was shown the suspect's photograph for five seconds. They were then asked to count backwards by threes for one minute. At the end of the minute, participants were asked to provide a description of the suspect based on the following instructions, "Imagine that the person you saw earlier committed a crime. Describe what she looked like, the way you would describe her to the police, if you wanted to help the police catch her." A composite description was developed by including any feature that was mentioned by more than half of the participants. The composite description was, "A young blonde woman."

To create the culprit absent lineup, I selected photographs of six individuals who matched this general description from a database maintained in my supervisor's lab. Photos were arranged in two rows of three. Each participant saw the photos in a randomly determined order. All photographs showed color head shots of the individuals and were cropped just above the top of the head and just below the bottom of the chin. The photographs were 200 pixels by 200 pixels. The individuals depicted were shown with neutral expressions. To create culprit present lineups, I took the culprit absent lineup, and replaced one of the fillers with a picture of the suspect. Because there were six fillers, I created six different culprit present lineups, with each culprit present lineup being created by replacing a different filler with the suspect.

To test the fairness of the lineups, I conducted a mock witness paradigm (Malpass & Lindsay, 1999). In a mock witness paradigm, participants who did not view the crime are provided with the composite description and are asked to try to select the suspect based on the description alone. If witnesses select the suspect more than what would be expected by chance (i.e., 1/6 of the time), the lineup is considered biased against the suspect. Seven-hundred and sixteen participants were recruited via Prolific and were paid \$0.25 for their participation. Each participant was randomly presented with one of the seven lineups. They were provided with the composite description and were asked to select the photograph that best matched the description. The proportion of times the suspect was chosen in the mock witness paradigm, across the six culprit present lineups, is shown in Table 1. For each lineup, I compared the proportion of times the suspect was selected to the proportion expected by chance (.1667) by means of a z test for proportions. Selection of the suspect did not significantly differ from chance for any of the lineups.

Table 1. Results of Mock Witness Paradigm for Culprit Present Lineups.

	Culprit Present 1	Culprit Present 2	Culprit Present 3	Culprit Present 4	Culprit Present 5	Culprit Present 6
Proportion Suspect Selections	0.13	0.17	0.13	0.14	0.16	0.18
n	102	102	102	102	103	103
z	1.06	0.00	1.06	0.80	-0.31	0.48
p	0.29	1.00	0.29	0.42	0.76	0.63

For my culprit absent lineup, there was no designated innocent suspect. For that reason, the fairness of the culprit absent lineup cannot be evaluated in terms of how often the suspect was selected. Instead, I determined the fairness of the culprit absent lineup by calculating Tredoux's E (Tredoux, 1998). Tredoux's E is based on the idea that, in a perfect lineup, each filler would be equally likely to be selected by a mock witness. Tredoux's E can range from 1.0

(extreme bias) to 6.0 (perfect fairness). Tredoux's E for the culprit absent lineup was 5.45 (95% CI: 4.97-6.03). Malpass (1981) argued that if the effective size of a lineup is at least 80% of its nominal size, then it can be considered fair. Nominal size refers to the actual number of people in the lineup. Tredoux's E is a measure of effective size. A Tredoux's E of 5.45 is 90.83% of the nominal size of 6. Based on the criterion suggested by Malpass (1981), it would be considered a fair lineup.

Procedure

Participants were presented with the experiment on the Qualtrics survey platform. After providing informed consent, participants were randomly assigned to view the brief duration video or the long duration video. I over-sampled the brief duration condition so that participants were 18 times more likely to be randomly assigned to the brief duration than the long duration condition. The settings on Qualtrics were set up so that participants could not advance the survey until either 12 seconds (brief duration condition) or 50 seconds had passed.

After viewing the film, participants were asked if they experienced any technical difficulties watching the film. If they said "Yes", they were asked to describe the nature of these technical problems. Typically, the technical difficulty involved excessive buffering. Participants then answered two attention check questions. One was a multiple-choice question asking for the color of the car. The correct answer was "black". The other asked the participant to select what was taken from the car. The correct answer was "purse".

Participants then completed a filler task to allow some time for the participant's memory to fade. The task was to solve as many simple arithmetic problems as they could in 5 minutes (e.g., $5+3-7 = ?$). Participants then moved on to the lineup task. A screen appeared showing them fair pre-lineup instructions (Lampinen et al., 2020). These are the types of instructions that

police provide witnesses in actual investigations. The instructions read, “Earlier you saw a woman take a purse from a car. We are now going to show you some pictures. If you see the perpetrator, please select her photo. If you do not see the perpetrator, select 'not present'. Please keep in mind the following: 1. the perpetrator may or may not be present in these photographs. 2. it is as important to clear an innocent person as it is to identify a guilty person 3. regardless of your choice, the police will continue to investigate the crime 4. after your decision, we will ask you to indicate how certain you are.”

After this, a randomly assigned culprit-absent or culprit-present lineup was administered to the participant. Participants had the option of selecting one of the six photographs or selecting a “Not Present” option. Regardless of their choice, participants were asked to evaluate their confidence in their identification choice, on a scale from 0% to 100%, in 10% increments. Following the confidence indication, participants were asked demographic questions. Participants were then debriefed and thanked for their time.

Results

The purpose of the present experiment is to analyze the effects of exposure duration on the confidence/accuracy relationship in an eyewitness identification task. In the present study, the suboptimal condition was the brief duration condition. Under this condition, I expected that highly confident individuals would not be as accurate as highly confident individuals in the long duration condition.

The mean amount of time spent on the task by participants in the brief duration condition was 4.93 minutes (SD = 2.77 minutes). Ninety-five percent of participants took between 3.28 minutes and 10.15 minutes. In the long duration condition, the mean amount of time spent on

the task was 6.03 minutes (SD = 8.15 minutes). Ninety-five percent of participants spent between 3.34 minutes and 13.91 minutes.

Accuracy Data

I performed chi-squared goodness of fit tests to determine if there was a significant difference in accuracy between the brief and long durations conditions. These comparisons were made separately for the culprit present and culprit absent lineups. When conducting these tests, participant responses were recorded as correct or incorrect and comparisons were made between the brief duration and long duration condition.

Table 2. Percentage of Suspect, Filler, and Reject Decisions.

	Suspect ID	Filler ID	Reject
Culprit Absent			
Brief Duration	n/a	46.3	53.7
Long Duration	n/a	29.8	70.2
Culprit Present			
Brief Duration	71.7	11.6	16.6
Long Duration	81.1	18.9	0.0

Table 2 shows the percentage of suspect identifications, filler identifications, and lineup rejections. For culprit absent lineups, there was no designated innocent suspect. For that reason, I only list filler identifications and lineup rejections for culprit absent lineups. For culprit present lineups, there was not a significant difference in the accuracy rate between the brief and long duration conditions, $\chi^2(df = 1, N = 991) = 2.20, p = .14$. For culprit absent lineups, mistaken identification of fillers was significantly more common in the brief duration condition than in the long duration condition, $\chi^2(df = 1, N = 988) = 5.88, p = .015$.

Confidence Data

The pristine conditions hypothesis says that when witnessing conditions are poor, there will be fewer high confidence identifications but the high confidence identifications that do

occur will still be highly accurate. In this thesis, poor viewing conditions involved having a brief view of the perpetrator's face and good viewing conditions involved having a longer view of the perpetrator's face. The pristine conditions hypothesis was examined by plotting confidence accuracy characteristics curves (CAC). These curves plot the accuracy of suspect identifications as a function of confidence. CAC curves specifically look at the accurate of suspect identifications. Because the culprit absent lineups did not have a designated innocent suspect, I estimated the innocent suspect identification rate by dividing the overall rate at which fillers were chosen in culprit absent lineups by six. Accuracy at each confidence level is then determined according to this equation:

$$Accuracy = \frac{CSID}{CSID + estMSID}$$

CSID stands for the proportion of correct suspect identifications from culprit present lineups. estMSID stands for the estimated mistaken suspect identification rate. It is determined by taking the proportion of people identified in the culprit absent lineup and dividing that proportion by six. The logic for doing this is, in a real police lineup one of the six people is always the suspect. In a culprit absent lineup used in an experiment, there is no reason to believe that any of the fillers is any more likely to be the police suspect than any of the others.

Figure 1 shows the CAC curve for the present experiment.

As is clear, witnesses who were 90-100% confident were very accurate regardless of whether they had a long or brief view of the face. To compare conditions using CAC curves, researchers usually generate 95% confidence intervals around each data point on the CAC curve and then look for non-overlapping confidence intervals. The ninety-five percent confidence intervals (shown with error bars) were determined using the bootstrapping technique recommended by

Seale-Carlisle and Mickes (2016). The 0-20% confidence interval and 30-40% confidence intervals of the long duration condition do not include error bars because there were no observations in those confidence intervals. Some of the confidence intervals are very large, reflecting the fact that there were a very small number of observations in that confidence range (i.e., the 50-60% confidence interval for the long duration condition).

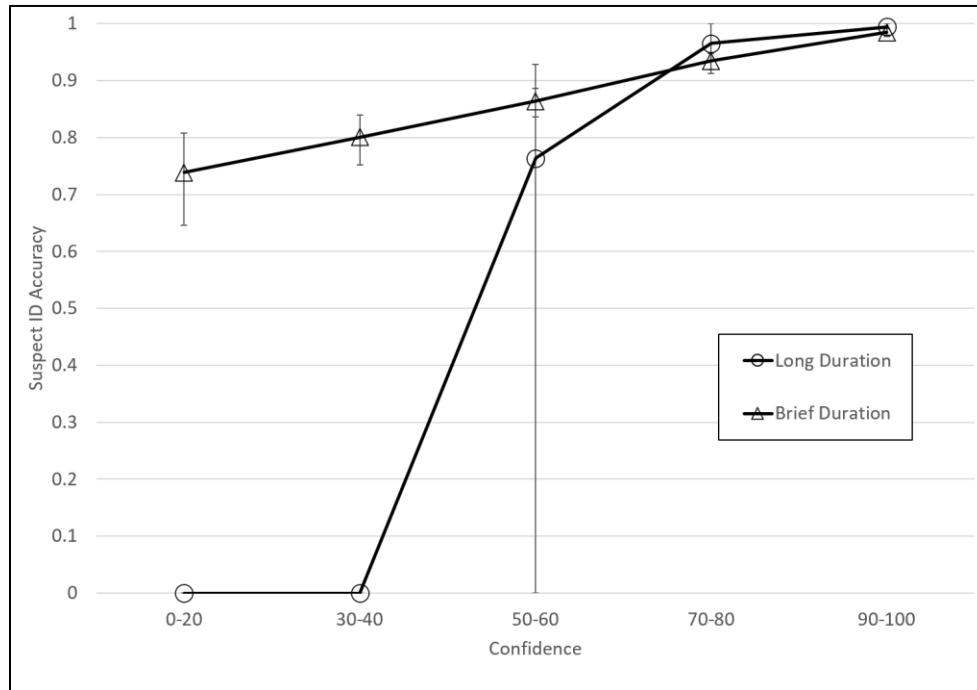


Figure 1. Confidence Accuracy Characteristic Curves.

For the purpose of evaluating the pristine conditions hypothesis, the most important confidence interval is the 90-100% confidence interval. Suspect identification accuracy in both conditions was very high when participants were 90-100% confident, regardless of how long they were exposed to the perpetrator's face. In the long duration condition, suspect identification accuracy was 99.4% when participants were 90-100% confident (CI: 98.04% - 1.00%). In the brief duration condition, suspect identification accuracy was 98.5% when participants were 90-100% confident (CI: 98.04-99.13). There was not a significant difference between the two conditions when it came to high confidence suspect accuracy. This is consistent with the pristine

conditions hypothesis (Wixted & Wells, 2017). The results indicate highly confident suspect identifications were highly accurate in both conditions. This finding was not what I hypothesized.

General Discussion

The pristine conditions hypothesis predicts that if system variables are optimal, eyewitnesses who are highly confident will be highly accurate, regardless of how poor the viewing conditions may be (Wixted & Wells, 2017). Contrary to the pristine conditions hypothesis, I predicted that if the viewing conditions of a crime are poor enough, then the eyewitnesses who are highly confident will not be highly accurate. I expected that if confidence was high, accuracy of suspect identifications would be high for the long duration condition but not the brief duration condition. These predictions were based on Giacona et al.'s (2021) prediction that when witnessing conditions get extremely poor, witnesses might under-correct the criterion they use to make their confidence judgments.

In the present study, I found that participants performed worse in the identification task in the brief duration condition than the long duration condition. This was expected due to the extremely short exposure of the face during the encoding process. However, the primary benefit of viewing the crime in the long condition, was that there were fewer mistaken identifications. It did not have a significant effect on the correct identifications. Participants were less likely to make high confidence suspect identifications when they were in the brief duration condition than when they were in the long duration condition. However, when participants in the brief duration condition were highly confident, they were highly accurate, similar to those in the long condition. Duration did not matter when participants were highly confident. This finding was

consistent with the pristine conditions hypothesis (Wixted & Wells, 2017) and disconfirmed my own hypothesis.

Discussion of Accuracy Results

The first result that will be discussed is that the memory of participants was better in the long duration condition than in the brief duration condition. This finding is consistent with previous research, and most people would probably assume this would occur (Gabbert et al., 2007). The effect of duration on memory is thought to reflect how visual attention works. When people look at an object, they sample information from that object over time. People's eyes move in quick jumps, called saccades, and are not focused on one single location in the object for a long time (Lampinen, Neuschatz, & Cling, 2012). Because the exposure to the face in the brief duration condition video was only 7 seconds long, there was less time for the eyes to move around the face, resulting in less information about the face being stored in memory.

Interpretation of Confidence Results

These results are also consistent with Semmler et al.'s (2018) claim that suboptimal estimator variables do not have an effect of the highly confident individuals' accuracy. Semmler et al. argued that participants are aware when viewing conditions are suboptimal, and this may cause a hesitancy in selecting a high level of confidence. This may explain why there were fewer high confidence identifications in the brief duration condition. Even though there were fewer indications of high confidence, Semmler et al. also concluded that when there were when witnesses were high in confidence, even in suboptimal viewing conditions, these participants would still be highly accurate.

On the other hand, there have been several recent studies that have found that even when pristine conditions are used, poor viewing conditions reduced the accuracy of highly confident

participants (Giacona et al. 2021; Lockamyeir et al. 2020; Ngyen et al 2017). One explanation for these contrasting results is that one needs to know exactly how a variable will affect memory in order to adjust the threshold for making a high confidence judgment (Giacona et al. 2021). For example, people would probably assume that when there is a weapon present in a crime, the memory of the crime will be improved because of the stress induced, when in actuality memory is worsened (Schmechel et al., 2005). If one does not understand how a variable might affect memory, one might under-correct for the impact of this variable, or correct in the wrong direction (Semmler et al., 2018).

Another explanation is that when people are attempting to correct their confidence judgements during extremely poor viewing conditions, they do not sufficiently correct for their viewing conditions. Consistent with this explanation, Lockamyeir et al., (2020) and Giacona et al., (2021) both found that high confidence did not necessarily imply high accuracy when viewing conditions were quite poor. Overall memory performance (as measured by d') was very poor in both experiments (Lockamyeir et al., 2020 & Giacona et al., 2021). In the present study, even participants in the brief duration condition performed relatively well. Thus, the reason I found evidence consistent with the pristine viewing conditions, might be because the poor viewing conditions were not sufficiently poor.

How the Present Study Compares to Prior Research

Memon et al. (2012) compared a long duration and brief duration condition. The long duration condition had an exposure of the culprit's face for about 45 seconds, and in the brief duration condition about 12 seconds. The advantage of viewing the long duration condition in Memon et al.'s study was strong. They found that young adults identified the culprit in the brief duration condition 25% of the time and 95% of the time in the long duration condition. When

young-adult participants were presented with a target-absent lineup, those who were in the brief duration condition made a mistaken identification 90% of the time, and this occurred 41% of the time in the long duration condition (Memon et al. 2003).

In the present study, I also found that the long duration condition produced better performance than the brief duration condition. However, the results were not quite as dramatic. I failed to find a significant difference between the brief and long duration conditions for culprit present lineups. For the culprit absent lineups, the proportion of mistaken identifications was significantly greater in the brief duration condition than the long duration condition, but the size of the difference was not as large as what was observed by Memon et al. (2003). This was despite the fact that my brief duration condition was shorter than Memon et al.'s brief duration condition.

One reason for the differing results might be that in Memon et al.'s (2002) study, the crime presented was more realistic, creating anxious or stressful feelings in the participants. The video used in Memon et al.'s (2003) was a professionally produced video re-enactment created by the British Broadcasting Corporation. The video used in my thesis involved amateur actors and was filmed by someone with no experience in cinematography. The crime depicted in the Memon et al. (2003) study was much more involved. It included a man who passed a note stating he had a gun and to fill the bag with money. There were other people present in the film, as well, making it more complex than the video used in my thesis.

In addition to these differences, Memon et al. (2003) used a filled retention interval of more than half an hour. This would have allowed a good deal of time for memory to decline. In my thesis, the retention interval was only 5 minutes long. Thus, in my thesis there was less time for memory for the perpetrator's face to decline. Given these differences, it is reassuring that the

basic finding replicates across the two experiments, even though it was stronger in Memon et al.'s experiment than in mine.

Limitations

It is interesting that high confidence suspect identifications were highly accurate regardless of viewing condition. This may be due to a number of limitations. One limitation is that the study lacked ecological validity in some ways due to the artificial nature of the crime and acting. It was very clear in the study that this was not a real crime and that no purse was really stolen. The acting was not done by someone with professional training. As a result, the intensity of the crime was not optimal, even for this being a non-violent crime.

Another limitation is that this study was all done on Prolific. Prolific pulls samples differently than if I were to have conducted an in-person study. The participants could have been intoxicated, sleep deprived, or in another suboptimal state when completing the study (Aruguete et al., 2019). Another concern is that the participants may have paused the video to get a better look at the face, if they suspected they were going to be tested on this. The participants may have also taken the study multiple times to be paid for multiple participations (Dennis et al., 2020). Although online data collection platforms like Prolific are a popular way to gain samples for research, there are several concerns about the collected sample. These participants derived from Prolific could be considered “professional experiment takers”, who participate in online experiments for money frequently.

Lastly, there is a possibility that the suspect in this present study had features that were too distinctive. This could be why participants had a seemingly easy time identifying the suspect, even in the brief duration condition, and were highly confident in their selections. As mentioned, a functional size analysis was conducted to ensure a fair lineup, so this would suggest that this

limitation is not a major concern. Although the functional size analysis was conducted and indicated that the lineup was fair, I do think there is a possibility the suspect stuck out among the fillers. Fillers were selected from a limited resource of photographs, so this could be why the difference in features between the suspect and fillers was arguably high.

Future Directions

If I were to continue research in this area, I would create a more realistic crime scenario to understand if this has an effect similar to Memon et al.'s (2003) findings. I would also test the same scenario in person, to see if there is a difference in performing this type of research online. I would make the threat of the crime a little more real, instead of something as harmless as stealing a stranger's purse.

I would also conduct the study with more diversity in the demographics to understand those differences related to the results. Because of the platform this study was conducted on, most participants were significantly older than the suspect of the mock crime they viewed. If done again, I would give this study to people closer in age to the suspect, along with people in various age ranges, to understand the effect this may have on eyewitness identifications and confidence.

This research joins other research in showing the brief exposure duration is associated with less accurate eyewitness memory. However, even though witnesses in the brief exposure condition were less accurate overall, the subset of identifications made with high confidence were very accurate (almost 100%). Undoubtedly, there is some limit to how briefly a witness can view a face, and still show the high confidence/ high accuracy pattern. But the present research provides the hopeful message that when police used recommended best practices, highly confident witnesses can often provide reliable evidence in the courtroom.

References

- Aruguete, M. S., Huynh, H., Browne, B. L., Jurs, B., Flint, E., & McCutcheon, L. E. (2019). How serious is the 'carelessness' problem on Mechanical Turk?. *International Journal of Social Research Methodology*, 22(5), 441-449.
- Bornstein, B. H., Deffenbacher, K. A., Penrod, S. D., & McGorty, E. K. (2012). Effects of exposure time and cognitive operations on facial identification accuracy: A meta-analysis of two variables associated with initial memory strength. *Psychology, Crime & Law*, 18(5), 473-490.
- Carlson, C. A., Dias, J. L., Weatherford, D. R., & Carlson, M. A. (2017). An investigation of the weapon focus effect and the confidence-accuracy relationship for eyewitness identification. *Journal of Applied Research in Memory and Cognition*, 6(1), 82-92.
- Clark, S. E., Howell, R. T., & Davey, S. L. (2008). Regularities in eyewitness identification. *Law and Human Behavior*, 32(3), 187-218.
- 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).
- Deffenbacher, K. A., Bornstein, B. H., McGorty, E. K., & Penrod, S. D. (2008). Forgetting the once-seen face: Estimating the strength of an eyewitness's memory representation. *Journal of Experimental Psychology: Applied*, 14(2), 139. doi:10.1037/1076-898x.14.2.139
- Dennis, S. A., Goodson, B. M., & Pearson, C. A. (2020). Online worker fraud and evolving threats to the integrity of MTurk data: A discussion of virtual private servers and the limitations of IP-based screening procedures. *Behavioral Research in Accounting*, 32(1), 119-134.
- Douglass, A. B., & Steblay, N. (2006). Memory distortion in eyewitnesses: A meta-analysis of the post-identification feedback effect. *Applied Cognitive Psychology*, 20(7), 859-869.
- Fawcett, J. M., Russell, E. J., Peace, K. A., & Christie, J. (2013). Of guns and geese: A meta-analytic review of the 'weapon focus' literature. *Psychology, Crime & Law*, 19(1), 35-66.
- Fitzgerald, R. J., Rubínová, E., & Juncu, S. (2021). Eyewitness identification around the world. In *Methods, measures, and theories in eyewitness identification tasks* (pp. 294-322). Routledge.
- Gabbert, F., Memon, A., & Wright, D. B. (2007). I saw it for longer than you: The relationship between perceived encoding duration and memory conformity. *Acta psychologica*, 124(3), 319-331.

- Giacona, A. M., Lampinen, J. M., & Anastasi, J. S. (2021). Estimator variables can matter even for high-confidence lineup identifications made under pristine conditions. *Law and Human Behavior*, 45(3), 256–270. <https://doi.org/10.1037/lhb0000381.supp> (Supplemental)
- Innocence Project (2009). Reevaluating Lineups: Why Witnesses Make Mistakes and How to Reduce the Chance of a Misidentification, <https://perma.cc/B5PKNU9A>
- Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence–accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(5), 1304 - 1316.
- Kovera, M. B., & Evelo, A. J. (2017). The case for double-blind lineup administration. *Psychology, Public Policy, and Law*, 23(4), 421.
- Lampinen, J. M., Erickson, W. B., Moore, K. N., & Hittson, A. (2014). Effects of distance on face recognition: Implications for eyewitness identification. *Psychonomic Bulletin & Review*, 21(6), 1489-1494.
- Lampinen, J. M., Neuschatz, J. S., & Cling, A. D. (2012). *The psychology of eyewitness identification*. Psychology Press.
- Lampinen, J. M., Smith, A. M., & Toglia, M. P. (2021). Eyewitness Memory: The Next 40 Years. In *Methods, Measures, and Theories in Eyewitness Identification Tasks* (pp. 377-389). Routledge.
- 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.
- Lockamy, R. F., Carlson, C. A., Jones, A. R., Carlson, M. A., & Weatherford, D. R. (2020). The effect of viewing distance on empirical discriminability and the confidence–accuracy relationship for eyewitness identification. *Applied Cognitive Psychology*. <https://doi.org/10.1002/acp.3683>
- Malpass, R. S., Tredoux, C. G., & McQuiston-Surrett, D. (2007). Lineup construction and lineup fairness. In *The Handbook of Eyewitness Psychology: Volume II* (pp. 169-192). Psychology Press.
- Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: Effects on eyewitness accuracy and confidence. *British Journal of Psychology*, 94(3), 339-354.
- 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.
- Nguyen, T. B., Pezdek, K., & Wixted, J. T. (2017). Evidence for a confidence–accuracy relationship in memory for same-and cross-race faces. *Quarterly Journal of Experimental Psychology*, 70(12), 2518-2534.
- Nyman, T. J., Antfolk, J., Lampinen, J. M., Tuomisto, M., Kaakinen, J. K., Korkman, J., & Santtila, P. (2019). A stab in the dark: The distance threshold of target identification in low light. *Cogent Psychology*, 6(1), 1632047.
- Police Executive Research Forum (2013). A National Survey of Eyewitness Identification Procedures in Law Enforcement Agencies. https://www.policeforum.org/assets/docs/Free_Online_Documents/Eyewitness_Identification/a%20national%20survey%20of%20eyewitness%20identification%20procedures%20in%20law%20enforcement%20agencies%202013.pdf
- Quigley-McBride, A., & Wells, G. L. (2021). Methodological considerations in eyewitness identification experiments. In *Methods, measures, and theories in eyewitness identification tasks* (pp. 85-112). Routledge.
- Seale-Carlisle, T. M., & Mickes, L. (2016). US line-ups outperform UK line-ups. *Royal Society Open Science*, 3(9), 1-9. <https://doi.org/10.1098/rsos.160300>
- Semmler, C., Dunn, J., Mickes, L., & Wixted, J. T. (2018). The role of estimator variables in eyewitness identification. *Journal of Experimental Psychology: Applied*, 24(3), 400-415.
- Sonenshein, D. A., & Nilon, R. (2010). Eyewitness Errors and Wrongful Convictions: Let's Give Science a Chance. *Or. L. Rev.*, 89, 263.
- Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, 36(12), 1546.
- 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.
- Wells, G. L., Rydell, S. M., & Seelau, E. P. (1993). The selection of distractors for eyewitness lineups. *Journal of Applied Psychology*, 78(5), 835.
- Wells, G. L., Steblay, N. K., & Dysart, J. E. (2015). Double-blind photo lineups using actual eyewitnesses: An experimental test of a sequential versus simultaneous lineup procedure. *Law and Human Behavior*, 39(1), 1.

Wixted, J. T., Read, J. D., & Lindsay, D. S. (2016). The effect of retention interval on the eyewitness identification confidence–accuracy relationship. *Journal of Applied Research in Memory and Cognition*, 5(2), 192-203.



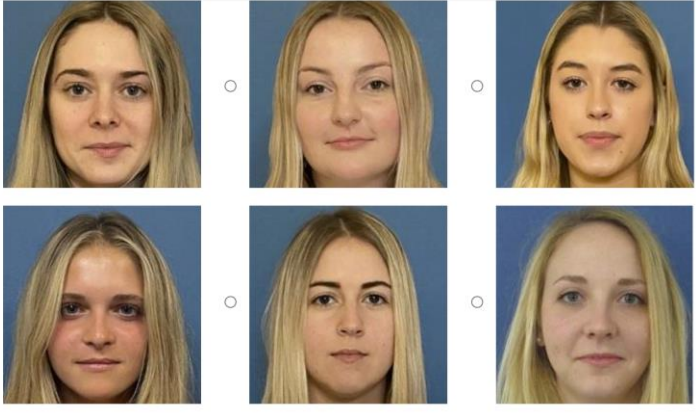
Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, 18(1), 10-65.















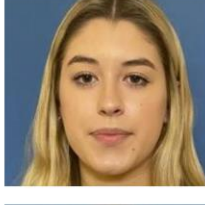



Appendix A. Demographic Characteristics of Sample.

Mean Age in Years	37.53
Standard Deviation Age in Years	14.26
Minimum Age in Years	18
Maximum Age in Years	81
Female	58.01%
Male	37.49%
Transgender Female	0.05%
Transgender Male	0.1%
Nonbinary	1.82%
Did not Reply	0.05%
Other	0.05%
Listed Age Rather Than Gender ₁	2.43%
Asian	10.16%
Black / African American	7.43%
Hispanic	5.1%
Native American	0.4%
White / Caucasian	71.4%
Biracial/Multiracial	4.7%
Other Replies ²	0.71%
Did not Reply	0.1%

Note: Participants were asked to indicate their gender and their race/ethnicity in their own words. Self-reported responses were then grouped into categories based on the judgment of the researcher. A small number of people indicated their age both for the question about age and the question about gender.¹ “Other Replies” for race/ethnicity included “American” and “Yes”.²

Appendix B Lineups Used in the Present Experiment

<p>Culprit Absent Lineup</p>	 <p>○ not present</p> <p>This lineup consists of six photographs of women with blonde hair, arranged in two rows of three. Each photograph has a small circle to its left, indicating a selection option. The top row shows three different women, and the bottom row shows three different women. The text 'not present' is located at the bottom left of the grid.</p>
<p>Culprit Present 1</p>	 <p>○ not present</p> <p>This lineup consists of six photographs of women with blonde hair, arranged in two rows of three. Each photograph has a small circle to its left, indicating a selection option. The top row shows three different women, and the bottom row shows three different women. The text 'not present' is located at the bottom left of the grid.</p>
<p>Culprit Present 2</p>	 <p>○ not present</p> <p>This lineup consists of six photographs of women with blonde hair, arranged in two rows of three. Each photograph has a small circle to its left, indicating a selection option. The top row shows three different women, and the bottom row shows three different women. The text 'not present' is located at the bottom left of the grid.</p>

<p>Culprit Present 3</p>	<p><input type="radio"/>  <input type="radio"/>  <input type="radio"/> </p> <p><input type="radio"/>  <input type="radio"/>  <input type="radio"/> </p> <p><input type="radio"/> not present</p>
<p>Culprit Present 4</p>	<p><input type="radio"/>  <input type="radio"/>  <input type="radio"/> </p> <p><input type="radio"/>  <input type="radio"/>  <input type="radio"/> </p> <p><input type="radio"/> not present</p>
<p>Culprit Present 5</p>	<p><input type="radio"/>  <input type="radio"/>  <input type="radio"/> </p> <p><input type="radio"/>  <input type="radio"/>  <input type="radio"/> </p> <p><input type="radio"/> not present</p>

<p>Culprit Present 6</p>	<table border="0"><tr><td data-bbox="636 281 652 302"><input type="radio"/></td><td data-bbox="662 201 854 390"></td><td data-bbox="870 281 886 302"><input type="radio"/></td><td data-bbox="896 201 1088 390"></td><td data-bbox="1104 281 1120 302"><input type="radio"/></td><td data-bbox="1130 201 1321 390"></td></tr><tr><td data-bbox="636 491 652 512"><input type="radio"/></td><td data-bbox="662 411 854 600"></td><td data-bbox="870 491 886 512"><input type="radio"/></td><td data-bbox="896 411 1088 600"></td><td data-bbox="1104 491 1120 512"><input type="radio"/></td><td data-bbox="1130 411 1321 600"></td></tr><tr><td data-bbox="636 617 652 638"><input type="radio"/></td><td colspan="5" data-bbox="662 617 1321 638">not present</td></tr></table>	<input type="radio"/>		<input type="radio"/>		<input type="radio"/>		<input type="radio"/>		<input type="radio"/>		<input type="radio"/>		<input type="radio"/>	not present				
<input type="radio"/>		<input type="radio"/>		<input type="radio"/>															
<input type="radio"/>		<input type="radio"/>		<input type="radio"/>															
<input type="radio"/>	not present																		

Note: For each participant, the ordering of the photos within the lineup was determined randomly.