5-2022

Diagnostic Featural Detection or Filler Siphoning: A Red Box Study

Brynn Schuetter
University of Arkansas, Fayetteville

Diagnostic Featural Detection or Filler Siphoning: A Red Box Study

An honors thesis submitted in partial fulfillment of the requirements of

Honors Studies in Psychology

By

Brynn Schuetter

Spring 2022

Psychology

J. William Fulbright College of Arts and Sciences

The University of Arkansas

Acknowledgments

I would like to thank both my thesis advisor, Dr. James Lampinen, and my graduate student advisor, Amber Giacona, for their help and guidance throughout the thesis process. I would also like to thank the research assistants in the LAMP lab who also helped throughout this process. Thanks are also due to Melissa Colloff and John Wixted for providing their materials. Finally, I would like to thank the members of my thesis committee, Dr. James Lampinen, Dr. Grant Drawve, Dr. Darya Zabelina, and Dr. Brandon Bouchillon for their time dedicated to this project.

Table of Contents

Introduction

According to the Innocence Project (2017), in nearly 70% of all DNA exoneration cases, mistaken eyewitness identification played a role in the wrongful conviction. Past research has shown that when a guilty suspect is present within a lineup, that guilty person is only selected from the lineup 46% of the time (Clark et al., 2008). The same meta-analysis found that when the suspect in the lineup is innocent, the witness still chooses someone from the lineup nearly 48% of the time. Another field study has found that one in every four witnesses identifies a known innocent person from a lineup, and nearly 40% of those who chose someone from a lineup, chose an innocent filler (Wells & Olson, 2003). Because of these faults, research within legal psychology has long been concerned with eyewitness identifications and how these procedures can be improved. There are a lot of factors that can affect the accuracy and reliability of an eyewitness and their identification. To collect an identification, an officer from a police department can give a witness a variety of identification procedures to complete; however, there are two main identification procedures. Lineups and showups are the most common identification procedures used by police in the United States.

The first type of identification procedure that officers can administer is called a lineup. Lineups can be administered either live or through photographs. In the U.S, a typical lineup contains one suspect and five fillers. The suspect is a person whom the police believe may have been involved in the crime. The fillers in a lineup are known to be innocent and could not have committed the crime. They provide protection to an innocent suspect by providing plausible alternatives to the suspect. The witness viewing the lineup is not aware of who the suspect is or their position in the lineup. In fact, a lineup is seen as unfair and biased if the witness is aware of who the suspect is within the lineup (Lindsay & Wells, 1980). Lineups can be administered

either sequentially or simultaneously. In a simultaneous lineup, all six of the photos are viewed at the same time, and one decision is made regarding who the perpetrator is. In a sequential lineup, photographs in the lineup are shown one at a time. The witness viewing the lineup is asked after each photograph if that person was the perpetrator of the crime. In both simultaneous and sequential lineups, the witness can either choose someone from the lineup as the perpetrator, or they can reject the lineup saying that the perpetrator was not present within the lineup.

The second type of identification procedure that police can give is called a showup. Like lineups, showups can either be live or through photographs. In a showup, the police show an eyewitness one person (i.e., the suspect) and then ask the witness if this person is the perpetrator of the crime that they witnessed. The only available answers to the witness are yes, this person is the perpetrator, or no, this person is not the perpetrator of the crime. This means that if a witness to a crime decides to choose someone out of the showup, the only option that they have is to choose the suspect. There is no protection available to the suspect like there is in lineup procedures.

Previous research has shown that lineups are superior to showups (Gronlund et al., 2012; Steblay et al., 2003; Yarmey et al., 1996). Although some research has shown that showup procedures tend to produce slightly more correct identifications, they were also shown to produce substantially more incorrect identifications when compared to lineups (Steblay et al., 2003). This means that in cases where the perpetrator is present within the showup, there is a higher hit rate, meaning that the perpetrator is correctly chosen from the lineup more often. This also means when the perpetrator is absent from the showup, there is a higher rate of false identifications as compared to lineups. Research has shown that showups are more prone to errors causing an increase in the number of mistaken identifications (Eisen et al., 2017; Steblay

et al., 2003). This effect was even more substantial when the suspect was presented in the showup wearing the same clothes that the perpetrator was described as wearing (Dysart et al., 2006; Yarmey et al., 1996). This has been named the clothing bias and is another reason why showups have been seen as inherently suggestive and biasing (Yarmey et al., 1996). Because of these errors, lineups have been shown to consistently produce more accurate and reliable identifications (Gronlund et al., 2012; Steblay et al., 2003; Yarmey et al., 1996).

One argument for the use of showups is that they can be used in situations where an identification needs to be made quickly. Research has shown that even in cases where showups are given immediately after witnessing a crime, they were never more accurate than a lineup was (Wetmore et al., 2015). The only advantage to showups is that they take fewer police resources than it takes to create a lineup. Fair lineups can be difficult to make and take time that is not always present in criminal cases. It can be difficult to find five fillers who resemble either the suspect, especially if the police want to administer a live lineup and have to coordinate the schedules of all five fillers and the suspect. Though lineups have long been shown to be superior to showups, showups are still widely used across the United States. A field study revealed that over 70% of identification tasks administered in the United States were showup tasks (Dysart & Lindsay, 2007).

Although it is widely agreed that lineups are more accurate and less suggestive than showups, there is less agreement about why lineups are superior (Wells et al., 2020). Two theories have been discussed as to why this may be the case. The first theory that has been proposed as to why lineups are superior to showups is called the diagnostic feature detection theory (Wixted & Mickes, 2014). This theory aims to explain why the fillers in a lineup add to the accuracy and reliability of the lineup procedure. This theory was developed to account for the

fact that simultaneous lineups have higher levels of empirical discriminability as compared to showups and other forms of identification procedures where faces are presented in isolation (Gronlund et al., 2012). The diagnostic feature detection theory claims that being able to compare members of the lineup to each other helps discriminate between guilty and innocent parties (Wixted & Mickes, 2014).

Within diagnostic feature detection theory, two types of features are defined: diagnostic features and nondiagnostic features. Nondiagnostic features are features that all fillers and the suspect should have in common. These commonalities should come from the witness' description of the perpetrator. For example, if the witness statement says that the perpetrator was a white male with dark hair, then all the members in the lineup should be white, a male, and have dark hair. These features are nondiagnostic. Diagnostic features are features that can differ from person to person within the lineup because they weren't mentioned in the witness's description of the perpetrator. These diagnostic features are the features that witnesses should use to differentiate between the guilty and innocent. Using the same perpetrator description, nondiagnostic features of those included in the lineup could be eye color, face shape, or facial markings because these features were not mentioned in the description of the perpetrator. Diagnostic feature detection theory states that by having the fillers in a lineup procedure, witnesses can use diagnostic features to differentiate between the guilty and innocent in the lineup before making a final decision. For example, if someone witnessed a robbery and the perpetrator had a black eye, but everyone present within the lineup had a black eye, the witness knows that this feature is not diagnostic and should not affect their final decision. This witness should use other features to determine who, if anyone, in the lineup is the perpetrator.

It has been shown that witnesses are more accurate when they notice both diagnostic and nondiagnostic information in procedures where photos are shown simultaneously instead of in isolation (Carlson et al., 2019). Previous research has shown that this is especially true in cases where the fillers are chosen because of their high similarity to the suspect. Researchers claim that having highly similar fillers allows for witnesses to better determine which features were most important in making an identification, therefore leading to more accurate identifications (Colloff et al., 2016; Colloff et al., 2018). If the fillers in a lineup are not fair, more features of the perpetrator will match with features of the suspect than with the fillers, meaning that more witnesses will be led to believe that the suspect is guilty. Past research has shown that as little as two fair fillers can increase the accuracy of a lineup (Wooten et al., 2020). Because there is only one person in a showup, and therefore no features to compare, witnesses may simply rely on the features given in their statement, which could lead to an inaccurate identification.

The second theory, filler siphoning theory, states that fillers draw choices away from the innocent suspect (Wells et al., 2015). In a showup, if the witness is going to pick someone, their only option is to pick the suspect. However, in a lineup, if the witness decides to pick someone, they have six people to choose from. This means that if the witness randomly picks someone, there is only a one in six chance that they will pick the suspect, but a five in six chance that they will pick someone who is known to be innocent. Filler siphoning theory also states that for the fillers to fulfill their purpose of adding protection to the suspect, the fillers must be plausible and competitive (Wells et al., 2015). The fillers within a lineup must all be similar to the suspect and match the general description that was given to the police by the witness. Previous researchers have claimed that the filler siphoning theory has even stronger evidence whenever the fillers are highly similar to the suspect because more highly similar fillers are moving decisions away from

innocent suspects (Smith et al., 2018). For example, when viewing a lineup, the witness has seven options that they could make. They could choose each of the members of the lineup, or they could reject the lineup as a whole. If the witness makes the choice to choose someone, filler siphoning theory states that there is only a one in six chance that the witness will choose the suspect should the suspect be innocent. If those fillers are highly similar to the suspect, there is a greater chance that one of the fillers will be chosen over the suspect. In an unfair lineup, filler siphoning theory states that the suspect will more closely resemble the witness's memory of the perpetrator, therefore increasing the likelihood that they will choose the suspect and a lower chance that they will be a filler.

Within the research community, there has been much debate over which of these two theories better explains why lineups are superior to showups. To test these theories, a novel identification procedure called the simultaneous showup was proposed (Colloff & Wixted, 2020). In this procedure, the experimenters gave the witness photos similar to what a witness would see in a simultaneous lineup procedure. Six photos were given to the witness, but the suspect's photograph was bordered in red. Participants were told that the person bordered in red was the police suspect and the other pictures were not suspects and were only included for comparison purposes. Participants were simply asked whether the police suspect (pictured in red) was the person who committed the crime. By creating this procedure, researchers were attempting to test the filler siphoning theory against the diagnostic feature detection theory. Colloff and Wixted (2020) argued that because the fillers were present, diagnostic feature detection could still occur. That is, even though the fillers could not be chosen, witnesses could still compare and contrast the photographs to see which features were diagnostic. Colloff and Wixted (2020) argued that filler siphoning was not possible, because the witnesses were unable

to choose any of the fillers. Results showed that participants in both the simultaneous showup and standard lineup conditions performed better than those in the standard showup condition (Colloff & Wixted, 2020). This was taken as evidence for the diagnostic feature detection theory because those in the simultaneous showup condition were not allowed to choose a filler, and therefore, arguably, filler siphoning could not happen.

Although the result seemed to support the diagnostic feature detection theory, these results could still be explained in a manner consistent with the filler siphoning theory. Even though participants were not able to choose one of the fillers in the simultaneous showup, they might have still concluded that one of the fillers was a better match to their memory than the suspect was (i.e., covert filler siphoning). For example, imagine a witness who views a simultaneous showup and decides that one of the fillers bears a striking resemblance to their memory of the perpetrator, while the suspect only bears a moderate resemblance. Presumably, in a lineup, the witness would pick the filler (filler siphoning). In a simultaneous showup, selecting the filler is not an option. But the presence of this highly similar (to their memory) filler would likely cause the witness to reject the less similar suspect. This reasoning leads to the conclusion that the simultaneous showup condition, does not rule out filler siphoning – because the filler siphoning might go on covertly, in the participant's mind.

The current study was interested in looking at the possibility that filler siphoning could still be present within the simultaneous showup condition. To examine this possibility, participants in the simultaneous showup condition were asked to make a decision and were then asked if the fillers in the task influenced their decision. If participants answered yes, they were then asked how those fillers influenced their decision. This allowed participants to admit that

they would have chosen someone else if given the option, therefore giving support to the filler siphoning theory.

The current study is a replication and extension of Experiment 3 conducted by Colloff and Wixted (2020). The aim of the current study was to further test the filler siphoning theory and the diagnostic feature detection theory by exploring how witnesses claim that the fillers from a simultaneous showup task affect their identification decisions. As the current study is a replication and extension of a previous study, the current study expects to reflect the results found by Colloff and Wixted (2020). It was also predicted that participants would indicate that the fillers influenced their decision when they indicated that the perpetrator was the suspect more often than when they indicated that the suspect was not the perpetrator. It was predicted that support for both filler siphoning and diagnostic feature detection theories would be given as explanations for how the fillers influenced identification decisions. Further, it was predicted that support for the filler siphoning theory would be given when participants indicated that the police suspect was not the perpetrator of the crime, and diagnostic feature detection support would be given when participants indicated that the police suspect was the perpetrator of the crime. Finally, it was also predicted that participants would show support for the filler siphoning theory when the perpetrator was absent from the simultaneous showup, and participants would provide support for the diagnostic feature detection theory when the perpetrator was present in the simultaneous showup.

Methods

Participants

There were 572 participants in the current study (M_{age}= 19.03, SD_{age}= 2.07). The majority of these participants were female (62.2%) and self-identified their race as White or Caucasian

(79.02%). All other demographic information can be found in Table 1. Participants were students at the University of Arkansas who were enrolled in a General Psychology course. They were recruited through SONA and given research credit for their participation in the study. Thirty-four participants were excluded from analysis because of misinformation, technical difficulties while completing the study, or incorrectly responding to the attention checks. The final sample size was 538.

Design

The study used a 3 (Presentation Style: simultaneous show-up, standard simultaneous lineup, standard showup) x 2 (Video: mugging, graffiti), x 2 (Culprit: present, absent) mixed design. Presentation Style was manipulated between subjects, while Culprit and Video were manipulated within subjects.

Materials

Materials in this study, including videos, lineups, and instructions were identical to those used by Colloff and Wixted (2020) and used with author permission.

Videos

We used two different 30-second-long videos. These videos were created by Colloff et al. (2016) and were also used in Colloff and Wixted (2020). Both videos were only visual, they did not contain any audio. One video depicts a white, male perpetrator taking a phone from another white male in a parking garage. For the first 20 seconds of the video, the victim is seen talking on the phone. The perpetrator then can be seen walking into the frame and saying a few words to the victim before stealing his phone. The perpetrator can be seen for nearly 10 seconds before he runs out of frame with the stolen property. The perpetrator in this video had a large scar on his cheek. The other video depicts a suspect with a black eye graffitiing a wall. The perpetrator's

face can clearly be seen for a little under 10 seconds before he turns around to spray paint the wall. For the remainder of the video, the participant can only see the back of the perpetrator, not his face.

Lineups and Showups

In the simultaneous lineup condition, a 3.5 inch by 3.5 black and white photo of the suspect (culprit-present) or filler (culprit-absent) was presented next to 5 other 3.5 inches by 3.5-inch black and white photos of innocent fillers. All fillers in the mugging condition were photoshopped to have the same scar that the perpetrator was seen to have. All fillers in the graffiti condition were photoshopped to have the same black eye that the perpetrator was seen to have.

In the standard showup condition, one 3.5 inches by 3.5 inches black and white photo of the culprit (culprit-present) or an innocent filler (culprit-absent) was presented. This photo was outlined in a red box to represent that they are the police suspect. If an innocent filler was presented, their picture was photoshopped to contain either the scar or the bruise according to their video condition.

In the simultaneous showup condition, the red box condition, a 3.5 inch by 3.5-inch black and white photo was presented with a red outline. This person was either the culprit (culprit-present) or an innocent filler (culprit-absent). The photo of the suspect was presented alongside five other 3.5 inches by 3.5-inch black and white photos of innocent fillers. These five photos did not have a red outline.

Procedure

The study was conducted online through Qualtrics. Participants in the study were first instructed to read and sign an informed consent form. After signing the consent form,

participants were told that they would be watching a video. They were instructed to pay careful attention, as they would have to answer questions about the video later. They were also told that they should only watch the video once. Participants were then randomly assigned to watch either the mugging video or the graffiti video first. After viewing the video, participants were asked if they experienced any technical problems while viewing the video, such as the video not loading or buffering. Participants were also asked two attention check questions regarding the contents of the video. These questions were easy to answer for anyone who had watched the video. If the participant viewed the mugging video, they were asked where the crime took place and what type of crime took place.

To create time between the viewing of the mock crime video and the completion of an identification task, participants were asked to complete a filler task. The filler task consisted of simple, 3-step math equations. For example, one of the equations could have been 8x6-2. Participants were asked to solve these math equations for 4 minutes.

After the filler task, participants were then given instructions on how to complete an identification task. They were randomly assigned to one of the three identification presentations (simultaneous show-up, standard simultaneous lineup, or standard show-up). The instructions varied depending on the condition that the participants were assigned.

In the simultaneous show-up condition, participants were told that they were going to be asked if the police suspect is the same man who was the perpetrator in the crime video that they just watched. They were told that the person whose photo was in the red box was the police's suspect and that "the other five men are not suspects; their role is to show you what an innocent suspect might look like in a case like this. The police suspect may or may not be the actual perpetrator." A simultaneous show-up would then appear, and participants were asked if the man

whose photo was highlighted in red was the man who committed the crime in the video. After completing the show-up, participants were asked if the photos of the five people not highlighted in red affected their decision. If participants answered that the other five people did affect their decision, they were asked to choose how those people affected their decision. Participants selected from the following options, with options randomized for each participant:

1. One or more of the other pictures (not in the red box) looked more like the perpetrator than the suspect (in the red box) did." (Filler Siphoning)
2.
3. "If I had been able to, I would have picked one of the pictures (not in the red box)". (Filler Siphoning)
4. "The other pictures (not in the red box) helped me to focus on what features were most relevant for the judgment." (Diagnostic Feature Detection)
5. "If the other pictures (not in the red box) were not there, I would have just focused on the fact that the suspect (in the red box) had a black eye [scar] like the perpetrator did." (Diagnostic Feature Detection)
6. "One or more of the other pictures (not in the red box) looked very similar to the suspect (in the red box) and this made me uncertain." (Filler Confusion)
7. "The other pictures (not in the red box) jogged my memory." (Reminding)
8. "The suspect (in the red box) looked more like the perpetrator than the men in the other photos (not in the red box)". (Filler Rejection)
9. "Other"

In addition to items meant to represent filler siphoning and diagnostic feature detection, we included other options that had been developed in a pilot experiment. One of these responses is filler confusion, in which one or more of the fillers looked so similar to the suspect that it made the witness unsure. Participants who indicated that the other photographs (not in red) influenced their decision, were allowed to choose as many items from this list as they liked.

In the standard simultaneous lineup condition, participants were told that they were going to view a lineup of six photos. Once the lineup was displayed, participants were told that the

lineup may or may not contain the perpetrator. They were then asked to select the person that they believe to be the perpetrator or select the "not present" option if they believe that the perpetrator was not present in the lineup. None of the photos in this condition were highlighted in red. There were no follow-up questions after completing the standard simultaneous lineup.

In the standard show-up condition, participants were informed that they would view a photo of the police suspect. The photo that they were shown was highlighted in red. They were informed that the suspect may or may not be the actual perpetrator. Finally, they were asked if the suspect highlighted in red was the person who committed the crime in the video. There were no follow-up questions after completing the standard show-up.

After completing their randomly assigned identification task, participants completed this procedure again with the video that they viewed the second time (mugging or graffiti). The presentation condition that the participant was assigned did not change for the second video. The same procedures used for the first video were continued into the second video. Once the identification task had been completed for both videos, participants were asked to answer demographic questions and were debriefed.

Results

The purpose of this experiment was to replicate and extend experiment 3 conducted by Colloff and Wixted (2020). We were interested in examining if filler siphoning was still mentally present within the simultaneous showup procedure. We hypothesized that support would be found for both filler siphoning theory and diagnostic feature detection theory. Further, we predicted that filler siphoning support would be more common when participants believed that the suspect was not the perpetrator and when the perpetrator was absent from the showup. This would make filler siphoning support found more commonly in showup rejections. We also

predicted that support for diagnostic feature detection theory would be more common in cases where the predicant indicated that the suspect was the perpetrator and when the perpetrator was present in the showup.

Replication of Colloff and Wixted (2020)

Colloff and Wixted (2020) analyzed their data using ROC analyses. They found that the ROC curves for lineups and simultaneous showups overlapped each other suggesting equal memory performance, and both dominated the ROC curve for showups. This suggests that memory performance was better for simultaneous showups and lineups than it was for showups. Table 2 shows the suspect, filler, and rejection rates for the three conditions. As has been found in previous research, showups had both the highest correct suspect ID rate and the highest mistaken suspect ID rate. Figure 1 shows the ROC analysis for the three conditions. As in Colloff and Wixted (2020), the ROC curves for the lineup and simultaneous showup are overlapping indicating equivalent memory performance. The ROC for lineups cuts off (is truncated) somewhat earlier than the ROC for the simultaneous showup. This indicates that the response criterion for simultaneous showups is somewhat laxer. Both of these ROC curves dominate the ROC curve for the showup, indicating that memory performance for the showup is worse than the other two conditions. These results replicate what was found by Colloff and Wixted (2020) and provide evidence that our participants were treating the task similarly to how their participants were.

Influence of the Other Pictures

The replication of Colloff and Wixted (2020) is consistent with the claim that our participants understood the task in a way that was similar to how Colloff and Wixted's participants understood the task. I now turn to how participants explained their decisions.

Do Participants Report the Other Pictures Influencing Them? Participants in the simultaneous lineup condition were asked whether the other pictures, not in the red box, influenced their decision. Overall, participants indicated that in 55.9% of cases in the simultaneous lineup condition indicated that they were influenced by the other pictures. Participants were significantly more likely to say that the other pictures influenced them when in the culprit absent condition (62.4%) than the culprit present condition (49.7%), $\chi^2(df = 1, N = 354) = 5.29$, $p = .02$. Participants were more likely to say that the other pictures influenced them when they rejected the suspect (63.3%) than when they accepted the suspect (48.3%), $\chi^2(df = 1, N = 354) = 8.14$, $p = .004$. Figure 2 shows that the participants reported the other pictures had the least influence in those cases where the participant viewed a culprit present simultaneous showup and they selected the suspect.

Is There Evidence of Covert Filler Siphoning? The simultaneous showup procedure was created to prevent filler siphoning from occurring, by making it impossible for participants to pick one of the fillers (Colloff & Wixted, 2020). The logic was that if filler siphoning cannot happen, but the simultaneous showup still outperforms the standard showup, it would be evidence that filler siphoning is not what explains the lineup/showup difference. The purpose of this thesis is to test the possibility that covert filler siphoning happens anyway. That is, although the witness cannot choose one of the fillers, one of the fillers can still be a better match to their memory than the innocent suspect, and this could lead to rejecting the suspect. There were two items on the questionnaire that represented filler siphoning:

1. "If I had been able to, I would have picked one of the pictures (not in the red box)".
2. "One or more of the other pictures (not in the red box) looked more like the perpetrator than the suspect (in the red box) did"

If either of these were selected, it was counted as evidence of covert filler siphoning. Covert filler siphoning was self-reported in 27.4% of all simultaneous showups and in 49.0% of all cases where the participant self-reported that the other pictures influenced their judgment. Covert filler siphoning was significantly more common for culprit absent lineups (43.9%) than culprit present lineups (11.6%), $\chi^2(df = 1, N = 354) = 46.47$, $p < .001$. Covert filler siphoning was quite rate when participants identified the suspect (5.7%) and significantly more common when the participant did not identify the suspect (48.3%), $\chi^2(df = 1, N = 354) = 80.66$, $p < .001$. Figure 3 shows that the participants reported engaging in covert filler siphoning least often when they made suspect identifications and most often when they did not identify the suspect.

Is There Evidence of Diagnostic Feature Detection? Diagnostic feature detection refers to a process in which the presence of other features focuses the witnesses' attention on those features that differ among the faces. Two items on the questionnaire were meant to capture diagnostic feature detection:

1. "The other pictures (not in the red box) helped me to focus on what features were most relevant for the judgment"

2. "If the other pictures (not in the red box) were not there, I would have just focused on the fact that the suspect (in the red box) had a black eye like the perpetrator did".

If either of these were selected, it was counted as evidence for diagnostic feature detection. Diagnostic feature detection was self-reported in 32.2% of all simultaneous showups and in 57.6% of cases where the participant indicated that the other pictures influenced their decision. Reports of diagnostic feature detection were approximately equally common for culprit present (30.9%) as for culprit absent (33.5%) simultaneous showups, $\chi^2(df = 1, N = 354) = 0.27$, $p = .60$.

Diagnostic feature detection was self-reported approximately as often when the suspect was rejected (29.4%) as when the suspect was selected (35.1%), $\chi^2(df = 1, N = 354) = 1.28$, $p = .26$.

Figure 4 shows that the proportion of self-reported diagnostic feature detection was highest when a suspect was mistakenly identified from a culprit absent simultaneous showup. However, it is important to keep in mind that the number of cases in which a suspect was identified in a culprit absent simultaneous showup was low, only 28 cases altogether. Nevertheless, diagnostic feature detection was more commonly reported in those cases than in cases when the suspect was rejected from a culprit absent simultaneous showup, $\chi^2(df = 1, N = 173) = 4.07$, $p = .04$, when a suspect was rejected from a culprit present simultaneous showup, $\chi^2(df = 1, N = 63) = 3.96$, $p = .047$, but not in cases where the suspect was correctly identified from a culprit present simultaneous showup, $\chi^2(df = 1, N = 174) = 3.27$, $p = .07$.

How Common is Filler Confusion? In the pilot testing that was used to develop the response alternatives, we discovered that some participants reported that there was another picture that looked so similar to the suspect that they were unsure if the suspect was guilty. We characterize this as 'filler confusion' because the response does not necessarily indicate that the participants' choice would have been drawn to the filler. We included one response alternative that corresponded to filler confusion:

> "One or more of the other pictures (not in the red box) looked very similar to the suspect (in the red box) and this made me uncertain."

Filler confusion was self-reported 15% of the time and in 26.8% of cases in which the participant indicated that the other pictures influenced their decision. Filler confusion was approximately equally common in culprit present (14.4%) and culprit absent (15.6%) simultaneous showups, $\chi^2(df = 1, N = 354) = 0.74$, $p = .39$. Filler confusion was also about equally common when the suspect was identified (13.8%) as when the suspect was not identified (16.1%), $\chi^2(df = 1, N =$

354) = 0.37, p = .54. Figure 5 shows the proportion of filler confusion responses. There did not appear to be any major differences as a function of either lineup type or suspect selection.

Other Responses. The main focus of this thesis is filler siphoning and diagnostic feature detection. I have also explored the related idea of filler confusion. However, participants also had other response options to choose from when they indicated that the other pictures influenced their choices. For 15.3% of cases, participants indicated that the other pictures 'jogged' their memory. In 16.4% of cases, participants indicated that the suspect looked more like their memory of the perpetrator than the fillers did. In 0.8% of cases, participants indicated that they had some other reason, not listed, for how the other pictures influenced their choice.

Discussion

The current study was interested in examining to what extent filler siphoning is still present, in a covert form, within the simultaneous showup procedure. Because this was a replication and extension of a previous study, we expected to replicate the results found by Colloff and Wixted (2020). We also expected that participants would indicate that the fillers influenced their identification decision. Further, we expected that participants would give support to both filler siphoning theory and diagnostic feature detection theory as to how the fillers influenced their decision. We hypothesized that support for filler siphoning would be found when participants indicated that the suspect was not the perpetrator and when the perpetrator was absent from the showup. For diagnostic feature detection theory, we hypothesized that participants would give support for this theory when they indicated that the suspect was the perpetrator and when the perpetrator was present in the showup.

Overall, the results supported our hypotheses. We found that both the simultaneous showup and simultaneous lineup provided more accurate and reliable identifications, therefore

replicating results found by Colloff and Wixted (2020). Results also showed that over half of the participants reported that pictures of the fillers influenced their decision. In almost half of these cases where participants reported that the fillers influenced their decisions, they stated covert filler siphoning to explain how they were influential. Diagnostic feature detection support was stated for how the fillers were influential in nearly 58% of cases where it was reported that the fillers influenced the identification decision. We also found that in cases where participants reported that the fillers influenced their decision, support for filler confusion was found around a quarter of the time and other participants gave another reason besides filler siphoning, filler confusion, or diagnostic feature detection.

Past research, and the results of this study, have shown that lineups are superior to showups (Gronlund et al., 2012; Steblay et al., 2003; Yarmey et al., 1996). There have been two theories proposed to try and explain why this is the case. Filler siphoning theory states that lineups provide added protection to the suspect within the lineup. Diagnostic feature detection theory state that the fillers in a lineup allow for the comparison of features to help witnesses differentiate between guilty and innocent parties.

The goal of the simultaneous lineup in the study done by Colloff and Wixted (2020), was to test the two theories by having fillers for comparison but not allowing participants to choose those fillers as the perpetrator. This was supposed to allow diagnostic feature detection, but not filler siphoning. Colloff and Wixted (2020) found that the simultaneous showup procedure outperformed the showup procedure. These results were replicated in the current study, meaning that participants treated the tasks in the current experiment in the same way that they treated the tasks in the study done by Colloff and Wixted (2020). Both studies found that the ROC curves for the simultaneous showup and the simultaneous lineup procedure dominated the standard

showup. These two procedures were similar, but the simultaneous showup procedure had more liberal choosing rates.

Colloff and Wixted (2020), however, interpreted their results of the simultaneous showup outperforming showups as support for the diagnostic feature detection theory. They reported that filler siphoning could not have accounted for their findings because the participants were not able to choose the fillers. The current study has found that this is not the case. In a quarter of all of the simultaneous showup procedures, and over half of the simultaneous showup procedures where participants said the fillers influenced their decision, participants reported that filler siphoning was still occurring. The two reasons that were selected to support filler siphoning were "one or more of the other pictures (not in the red box) looked more like the perpetrator than the suspect (in the red box) did" and "if I had been able to, I would have picked one of the pictures (not in the red box)". This means that even though participants were not able to choose the fillers from the showup, they would have if they were given the chance. This thought causes them to say that the perpetrator was not the suspect in the red box.

For example, a participant could have viewed the simultaneous showup and thought that the second filler was the perpetrator. Because their only options are between if the suspect is the perpetrator or not, they are forced to answer that the perpetrator is not the suspect. The participant could have thought that both the second and third fillers looked more similar to the actual perpetrator. This too would force them to say that the suspect was not the perpetrator. In both of these cases, if the witness had been given a lineup, they would have chosen someone else within the lineup. This means that filler siphoning is still occurring in the simultaneous showup procedure, but it is covert because the option to choose a filler is taken away from them.

These examples can also be used to explain why filler siphoning options were chosen most often when a participant correctly rejected the showup. In these cases, the participant indicated that the suspect was not the actual perpetrator. These showups were target absent meaning that the perpetrator was not in the showup, so the participants made the correct decision in saying that the suspect was not the perpetrator. If one of the other fillers looked more like the actual perpetrator, then the only option for the participant was to choose that the suspect was not the perpetrator. In a standard simultaneous lineup, these participants would have chosen a filler and therefore giving support to the filler siphoning theory.

One criticism of eyewitness identification research is that the results might not generalize to the real world. Ethically, researchers cannot force participants to watch real crimes and ask them to pick the perpetrator from a lineup. Because of this, many eyewitness identification studies use mock crime videos similar to the videos used in the current study. There are several differences between watching a mock crime video and being a witness to an actual crime. There is also a significant amount of pressure put on real witnesses while viewing a lineup. A witness might believe that a guilty person will walk free if they do not pick the correct person from the lineup. They could also fear that they will choose an innocent person and be a factor in putting an innocent person behind bars. The participants in the current study likely did not feel any of these pressures and may not have taken the identification task as seriously as they would have in the real world. Because of this, we do not know if the behaviors of participants in the study would generalize to real-world identification tasks.

In order to increase the likelihood that a participant would take the study seriously and want to make the correct identification decision, future studies could incentivize the participants. The current study gave students course credit, but that does not incentivize participants to make

correct decisions. If for example, there was a monetary reward for correct answers, participants may be more likely to try their best in the identification task. This would also create a more realistic study because witnesses in the real world are motivated to make the correct decision.

The current research was interested in further examining a novel procedure called the simultaneous showup procedure to explore the relationship between lineups and two competing theories. In contrast to previous research discounting the filler siphoning theory and finding support for the diagnostic feature detection theory, the current project found support for both theories. Results showed that mock witnesses gave reasonings as to how photos of fillers influenced their decision that supports both diagnostic feature detection theory. Therefore, the current study has shown that covert filler siphoning does occur even when the option to pick a filler is taken away from the witness.

References

Carlson, C. A., Jones, A. R., Whittington, J. E., Lockamyeir, R. F., Carlson, M. A., & Wooten, A. R. (2019). Lineup fairness: Propitious heterogeneity and the diagnostic feature-detection hypothesis. *Cognitive research: principles and implications*, *4*(1), 1-16.

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., Wade, K. A., & Strange, D. (2016). Unfair lineups make witnesses more likely to confuse innocent and guilty suspects. *Psychological Science*, *27*(9), 1227-1239.

Colloff, M. F., Wade, K. A., Strange, D., & Wixted, J. T. (2018). Filler-Siphoning theory does not predict the effect of lineup fairness on the ability to discriminate innocent from guilty suspects: Reply to Smith, Wells, Smalarz, and Lampinen (2018). *Psychological Science*, *29*(9), 1552-1557.

Colloff, M. F., & Wixted, J. T. (2020). Why are lineups better than showups? A test of the filler siphoning and enhanced discriminability accounts. *Journal of Experimental Psychology: Applied*, *26*(1), 124.

Dysart, J. E., & Lindsay, R. C. L. (2001). A preidentification questioning effect: Serendipitously increasing correct rejections. *Law and Human Behavior*, *25*(2), 155-165.

Dysart, J. E., & Lindsay, R. C. L. (2007). Show-up identifications: Suggestive technique or reliable method?. In *The Handbook of Eyewitness Psychology: Volume II* (pp. 151-168). Psychology Press.

Dysart, J. E., Lindsay, R. C., & Dupuis, P. R. (2006). Show-ups: The critical issue of clothing bias. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *20*(8), 1009-1023.

Eisen, M. L., Smith, A. M., Olaguez, A. P., & Skerritt-Perta, A. S. (2017). An examination of showups conducted by law enforcement using a field-simulation paradigm. *Psychology, Public Policy, and Law*, *23*(1), 1.

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

Innocence Project (2017). *Eyewitness identification reform.* https://innocenceproject.org/eyewitness-identification-reform/

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

Kneller, W., Memon, A., & Stevenage, S. (2001). Simultaneous and sequential lineups: Decision processes of accurate and inaccurate eyewitnesses. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *15*(6), 659-671.

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

Lindsay, R. C., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology*, *70*(3), 556.

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

Smith, A. M., Wells, G. L., Smalarz, L., Lampinen, J. M. (2018). Increasing the similarity of lineup fillers to the suspect improves the applied value of lineups without improving memory performance: Commentary on Colloff, Wade, and Strange (2016). Psychological Science, 29, 1548–1551.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2003). Eyewitness accuracy rates in police showup and lineup presentations: A meta-analytic comparison. *Law and Human Behavior*, *27*(5), 523-540.

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.

Wells, G. L., & Olson, E. (2003). Eyewitness identification. *Annual Review of Psychology*, *54*(1), 277-295.

Wells, G. L., Smalarz, L., & Smith, A. M. (2015). ROC analysis of lineups does not measure underlying discriminability and has limited value. *Journal of Applied Research in Memory and Cognition*, *4*(4), 313-317.

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.

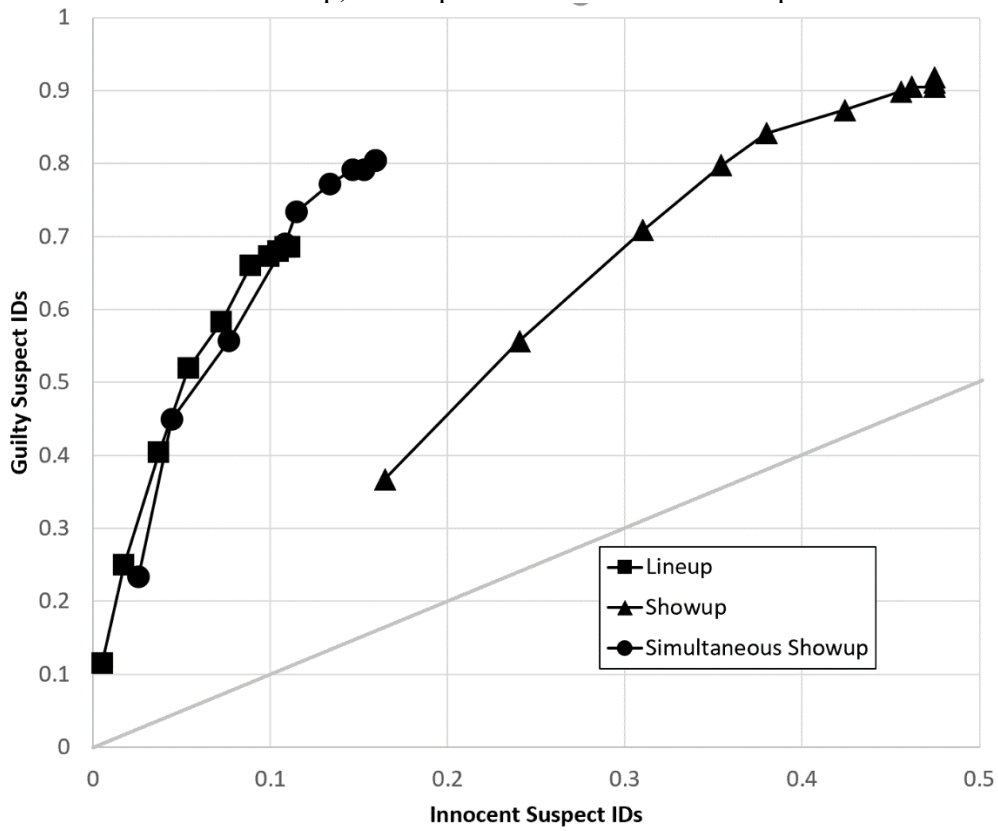
Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychological Review*, *121*(2), 262.

Wooten, A. R., Carlson, C. A., Lockamyeir, R. F., Carlson, M. A., Jones, A. R., Dias, J. L., & Hemby, J. A. (2020). The number of fillers may not matter as long as they all match the description: The effect of simultaneous lineup size on eyewitness identification. *Applied Cognitive Psychology*, *34*(3), 590-604.

Yarmey, A. D., Yarmey, M. J., & Yarmey, A. L. (1996). Accuracy of eyewitness identifications in showups and lineups. *Law and Human Behavior*, *20*(4), 459–477. https://doi.org/10.1007/bf01498981

Tables and Figures

Figure 1. ROC curves for the lineup, showup and simultaneous showup conditions.



Note: Grey line indicates chance performance.

Figure 2. Proportion of participants who indicated that the other pictures influenced their choice.

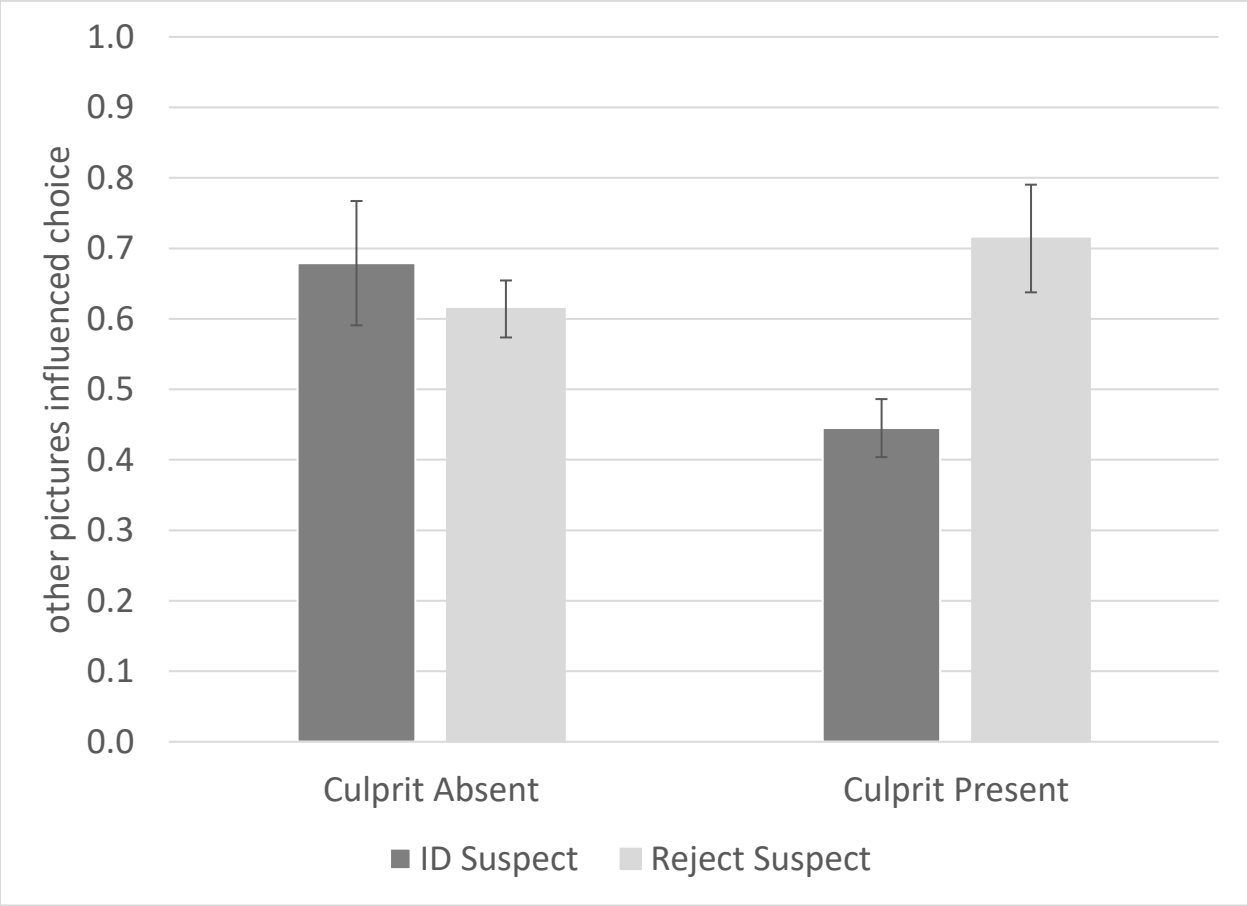


Figure 3. Proportion of participants reporting covert filler siphoning.

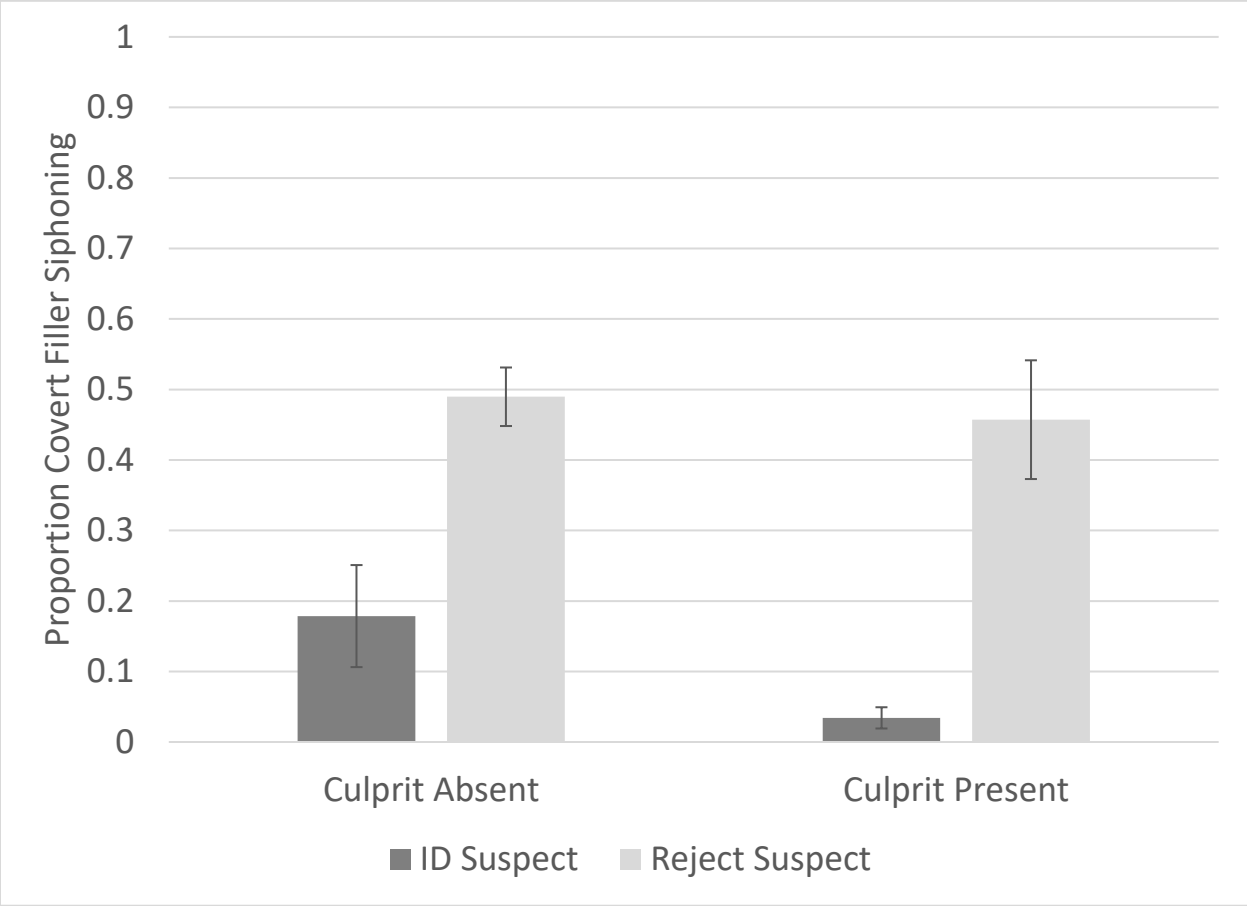


Figure 4. Proportion of self-reported diagnostic feature detection.

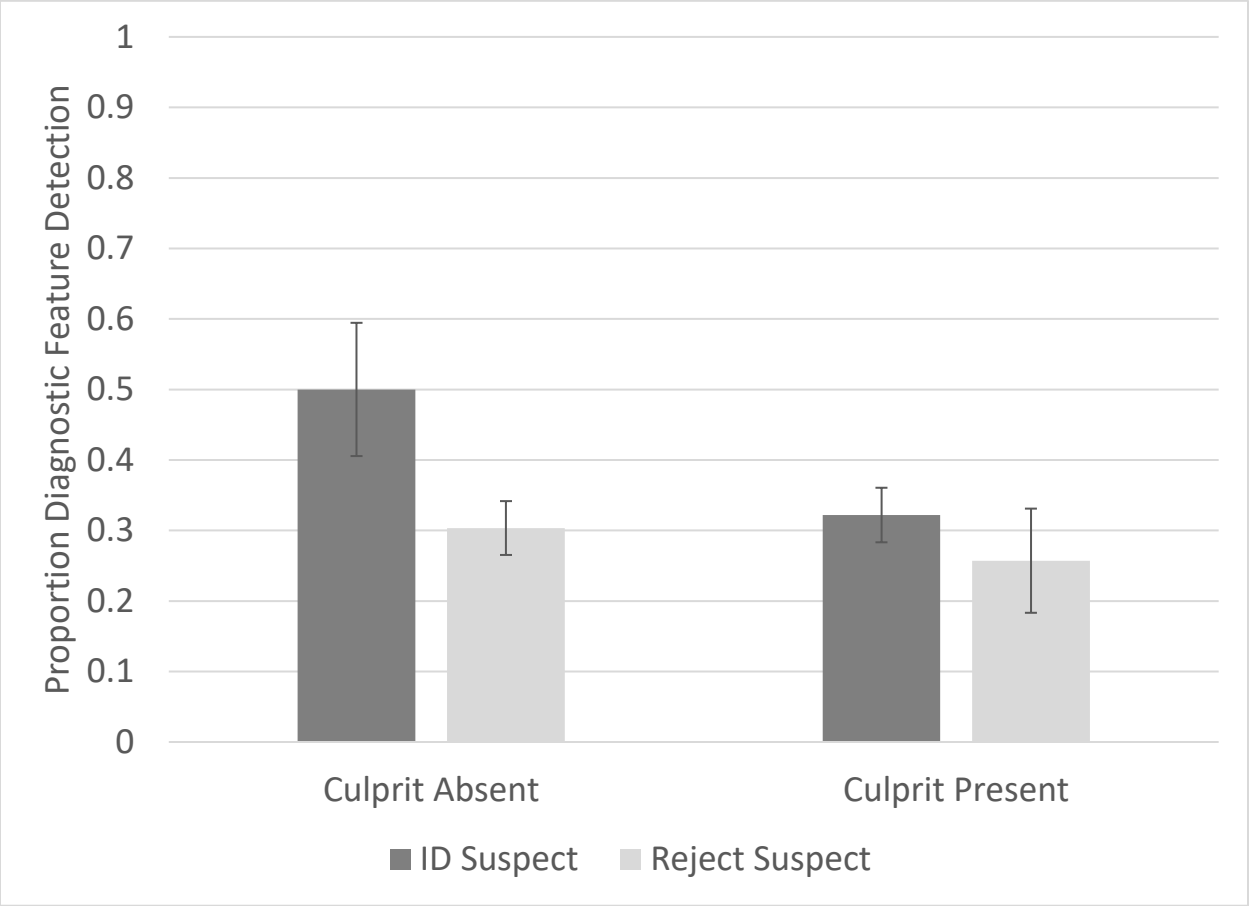


Figure 5. Proportion of filler confusion responses.

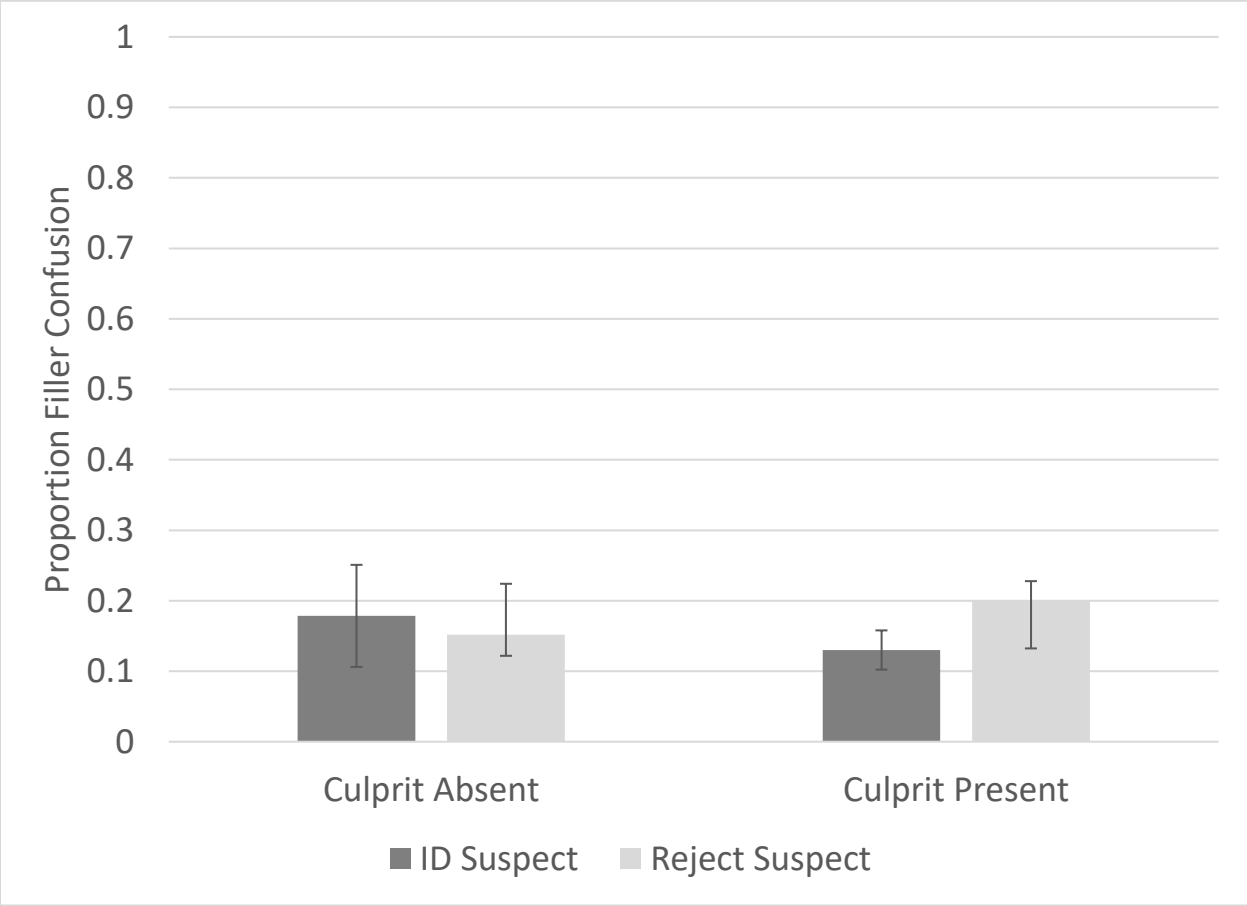


Table 1. Demographic Information.

Gender	
Female	356
Male	196
Other	3
Did Not Reply	17
Race/Ethnicity	
Asian	9
Black/African American	26
Native Hawaiian or Other Pacific Islander	1
Hispanic/Latinx	30
Middle Eastern	1
Native American	4
White/Caucasian	452
Multi-Racial/Ethnic	32
Did Not Reply	17
Age	
Mean Age (Years)	19.03
Standard Deviation (Years)	2.07
Minimum Age (Years)	18
Maximum Age (Years)	51

Note: For Gender and Race/Ethnicity, values represent frequencies out of 572 total participants.

Table 2. Percentage of suspect identifications, filler identifications and rejections as a function of condition.

	Suspect	Filler	Reject
Culprit Absent			
Lineup	10.7*	53.5*	35.8
Showup	48.6	n/a	51.4
Simultaneous Showup	17.6	n/a	82.4
Culprit Present			
Lineup	67.8	19.1	13.1
Showup	91.4	n/a	8.6
Simultaneous Showup	80.4	n/a	19.6

Note: Filler identifications are not possible in the showup or simultaneous showup and are marked a n/a for 'not applicable'. Suspect identifications in the culprit absent lineup is estimated by taking overall choosing rate and dividing by 6. Filler identifications in culprit absent lineup is estimated by taking overall choosing rate and multiplying by 5/6. Estimated values are indicated with *