Do the Clothes Make the Criminal? The Influence of Clothing Match on Identification Accuracy in Showups

Stacy A. Wetmore[1], Jeffrey S. Neuschatz[2], Scott D. Gronlund[1], Kylie N. Key[2], and Charles A. Goodsell[3]

[1]University of Oklahoma

[2]The University of Alabama in Huntsville

[3]Canisius College

Please address editorial correspondence to:

Stacy A. Wetmore
University of Oklahoma
Department of Psychology
455 W. Lindsey Street
Dale Hall Tower, Room 705
Norman, OK 73019-2007
Email: stacy.wetmore@ou.edu

Author Notes

This work was supported by the National Science Foundation (NSF) grants SES-1060913 to CAG, SES-1060921 to JSN, and SES-1060902 to SDG. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of the NSF.

Abstract

Showups, a single suspect identification, are thought to be a more suggestive procedure than traditional lineups by the U.S. Supreme Court and social science researchers. Previous research typically finds that a clothing match in showup identifications increases false identifications. However, these experiments do not allow for a determination of whether this increase arises from a change in response bias, reduced discriminability, or both. In the present study, participants viewed a mock crime video and made a showup identification with either a clothing match or mismatch. Contrary to prior research, the best discriminability occurred when the guilty and innocent suspects wore clothing that matched the clothing worn during the crime. A clothing match also resulted in a more liberal response bias. The results are consistent with the principle of encoding specificity and the outshining hypothesis, as instantiated in the Item, Context, Ensemble theory. Practical implications are discussed.

Keywords: Showups, Clothing Match, Receiver Operating Characteristic (ROC) analysis

Do the Clothes Make the Criminal? The Influence of Clothing Match on Identification Accuracy in Showups

Showups are a one-person identification procedure that is usually conducted soon after a crime has occurred, either in person or by presenting a photograph (see Goodsell, Wetmore, Neuschatz & Gronlund, 2013). The U.S. Supreme Court has declared that showups are an unreliable identification procedure (*Stovall v. Denno*, 1967) and should only be used in a limited set of circumstances. Nonetheless, showups are one of the most frequently used identification procedures; some researchers estimate that as many as 77% of cases with eyewitness evidence involve showups rather than lineups (Dysart & Lindsay, 2007). Given the frequency with which showups are conducted, research regarding the factors that affect the reliability of showups is of great interest. The focus of the current paper is on one factor that may influence the reliability of showups, clothing match.

Eyewitness descriptions often contain descriptions of clothing, which means that clothing match could play a role at identification (Dysart, Lindsay & Dupuis, 2006; Susa & Meissner, 2008). A clothing match, often referred to as a clothing bias in psycho-legal research, has generally produced an increase in false identifications. False identifications occur when an innocent suspect is chosen from a perpetrator absent identification procedure. For a lineup identification, a clothing match would make an innocent suspect stand out, which increases the likelihood that he will be chosen if the fillers – known innocents presented with the suspect – are not wearing the same thing (Dysart & Lindsay, 2007; Lindsay, Wallbridge, & Drennan, 1987). However, a clothing match could be even more problematic for showups. Individuals who are found near the crime scene, and happen to be wearing clothing similar to the perpetrator, are likely to be subjected to a showup identification. To the extent that an eyewitness bases a

showup identification decision in part on what a suspect is wearing, a clothing match could increase false identifications (Schmechel, O'Toole, Easterly, & Loftus, 2006). But it is important to note that a clothing match likely does not only affect false identification rates. To fully understand the impact of clothing match, we also must examine the rate of correct identifications (when a guilty suspect is chosen from a perpetrator present identification procedure).

Clothing match can affect performance in two ways: 1) by altering discriminability, the ability to differentiate guilty suspects from innocent suspects, or 2) through a response bias, the greater willingness of participants to choose from one condition than another. For instance, a clothing match may provide additional cues to aid memory for the perpetrator, but fail to increase the degree to which an innocent suspect resembles the perpetrator. This increase in correct identifications, coupled with a decrease (or lack of increase) in false identifications, would produce an increase in discriminability. Alternatively, if the clothing match entices participants to select more from a showup, both correct and false identifications may increase, indicating a shift to a more liberal response bias. Of course, it is also possible for there to be an increase to both correct and false identifications, but the increase for correct identifications could be greater than for false identifications. This would result in a change to discriminability and response bias.

Clothing and Context Match

A few studies have found a negative effect on false identifications when clothing is matched in showups. Yarmey, Yarmey, and Yarmey (1996) found that an innocent suspect – similar to the perpetrator in physical appearance (e.g., hair style and length, weight, etc.) – was falsely identified more often from showups when wearing the same clothing at identification as at the event (white T-shirt). Dysart, Lindsay, and Dupuis (2006) found a similar effect on false

identifications, but only when distinctive clothing (a Harley Davidson logo T-shirt, no effect for common clothing) was worn. However, there was no significant change in the correct identification rate as a function of clothing match versus mismatch in either study. Consequently, these results are consistent with an adverse effect of clothing match, an increase in false identifications, which produces an adverse effect on discriminability.

However, Valentine, Davis, Memon, and Roberts (2012) found no effect of clothing match. Although false identifications were higher when an innocent bystander was wearing a distinctive shirt during the incident, the increase was not significant. Correct identifications could not be evaluated because only perpetrator absent showups were conducted. Lawson and Dysart (2012) also reported no influence of a clothing match on own-race versus cross-race identification accuracy, but found that presenting the suspects in the same clothing worn by the perpetrator increased correct identifications by 14.6%. Unfortunately, false identification rates were not reported and could not be computed. Thus, it is unclear if the increase in correct identifications in the Lawson and Dysart study represents an increase in discriminability or a response bias.

Clothing match can be interpreted as a type of context, and the impact of context on retrieval has been addressed in the psycho-legal literature (Cutler, Penrod, O'Rourke, & Martens, 1986; Cutler, Penrod & Martens, 1987; Smith, Leach, & Cutler, 2012). Smith et al. (2012) examined whether matching the context of the event at a subsequent showup identification could enhance performance. The researchers varied whether the presence of the stolen property in question (a backpack) could provide an additional context cue to enhance memory retrieval. The results indicated that participants made more correct identifications when the backpack was present during the event and the identification. Furthermore, there was a non-significant decrease

in false identifications when the backpack was presented with the innocent suspect. The authors argued that there was an increase in discriminability when the backpack was present during the showup, without affecting response bias. However, it is important to note that in the stolen property condition there was a 100% identification rate from the perpetrator present showup. This ceiling effect makes it difficult to determine if the results were due to the context match or an idiosyncrasy of the stimuli.

In sum, the few studies that have addressed clothing and context match in showups have led to ambiguous findings for a variety of reasons. Yarmey et al. (1996) found increases in false identifications when the perpetrator wore common clothing. Neither Dysart et al. (2006) nor Valentine et al. (2012) replicated this effect with common clothing. The one study that evaluated context match, Smith et al. (2012), was difficult to interpret due to the ceiling effect. The ambiguous findings are surprising given that basic memory theory makes clear predictions about the effects of context match on memory performance.

Memory Theory and Clothing Match

At the most basic level, the showup is a recognition task and clothing match is a context effect. The encoding specificity principle predicts that the best memory performance (i.e., best discriminability) should arise when the cues available at encoding match the cues available at retrieval (Tulving & Thomson, 1973). However, as Nairne (2002) pointed out, memory performance is not based on the absolute match between cues and memory traces but on the relative degree of match. The relative match is the degree to which a cue compound (items integrated together into a whole) is distinctive and focuses on a target memory (Watkins & Watkins, 1975, 1976, called this cue-overload). Therefore, the discriminability of a guilty from an innocent suspect will be easier when there is a clothing match than when there is not a

clothing match because a compound cue of the face and clothing together targets the guilty suspect better than does a face or clothing cue alone. A compound cue that includes the innocent suspect's face will match memory more poorly, irrespective of the clothing match.

Alternatively, the outshining hypothesis (Smith, 1988; 1994) also may provide an explanation for the effect of clothing match on showup identification. Smith argued that strong retrieval cues "outshine" weaker cues. In his evaluation of recognition memory for words, he found that memory was not enhanced by a context cue if the word was a sufficiently strong cue (the word outshines the context). Therefore, if an eyewitness was able to adequately view a perpetrator (e.g., long exposure time, good lighting, close proximity, high visibility of features), then the face should be well encoded and the addition of context information (clothing match) would not change performance. However, actual crimes typically occur in less than ideal viewing conditions (e.g., poor illumination, weapon focus, stress, disguise). Consequently, if the face is encoded poorly then the clothing becomes a relatively stronger cue. Because the clothing is common to the perpetrator and innocent suspect, this should increase both correct and false identification rates, which is akin to a shift in response bias, but produces no change in discriminability.

Lastly, Murnane, Phelps and Malmberg (1999) proposed that recognition is based on three types of information: item, associated context, and ensemble (ICE). Murnane et al. define the item as the central focus of the primary cognitive task, the associated context is any information in the processing environment peripheral to the cognitive task, and the ensemble is the integration of the item and associated context. In terms of a showup, the item is the perpetrator's face, the associated context is the clothing, and the ensemble would be the unique integration of the face and the clothing. The encoding of an ensemble is optional, and likely is

more effortful. In fact, if elaborative encoding is required to create an ensemble, the stressful situation that surrounds a crime might make it difficult to construct.

The ICE theory subsumes encoding specificity and the outshining hypothesis. According to the theory, both correct and false identification rates will increase when the associated context (the clothing) is used as a cue, making the identification of both the perpetrator and the innocent suspect more likely (due to a shift in response bias). These predictions are consistent with the outshining hypothesis. However, if ensemble information is encoded and clothing is part of that compound cue, correct identifications should increase more than false identifications because the ensemble includes the perpetrator's face and not the innocent suspect. This is consistent with the predictions of encoding specificity.

Hypotheses

In order to evaluate the effect of clothing match on identification accuracy in showups, we varied two factors, shirt (same or different) and hat (match or mismatch). If the item information (face) is well encoded, there should be little effect of clothing match on discriminability or response bias as the face will "outshine" the clothing. However, if the face is not well encoded, the ICE theory can predict a difference in discriminability, response bias, or both, depending on how the item, associated context, and ensemble information are encoded. If clothing (shirt, hat) is encoded as associated context, it should result in increased choosing (a liberal response bias) and increased correct and false identifications in the match conditions (clothing outshines the face). However, if the shirt and hat are encoded as an ensemble with the face, it predicts a discriminability difference (due to encoding specificity). The reliance on the ensemble should produce an increase in correct identifications when the items in the event are identical to those presented at the showup identification. That is, participants should demonstrate

the best discriminability when the hat match and same shirt are provided at the identification. However, if the shirt and hat contribute to retrieval as *both* associated context and as an ensemble, a difference in discriminability and a shift to a more liberal response bias should arise.

Method

Participants

A total of (*N* = 2611) participants were recruited from the University of Oklahoma (*n* = 137), UAHuntsville (*n* = 94), Texas A&M University at Commerce (*n* = 216), Canisius College (*n* = 155), and the community (*n* = 1936). An additional 73 participants were excluded from data analysis for failure to correctly answer two designated manipulation-check questions. College participants received course credit; community participants earned points in accordance with the policy of the company providing the sample. There were 1518 men and 1016 women with a mean age of 38.7 years. Of those participants who indicated their ethnic background (*n* = 1571), most participants were Caucasian (76.1%), followed by African American (9.2%), Latin/Hispanic (6.7%), Asian/Pacific Islander (3.9%), Native American Indian (1.9%), and Other (2.2%). For those participants who indicated their educational background (*n*=1571), a majority indicated some college (45.3%), followed by high school/GED (23.1%), college graduate (22.9%), professional degree (7.9%), and Other (.8%). All guidelines set forth by the American Psychological Association and the Institutional Review Boards at the appropriate educational institutions were followed.

Design

The experiment utilized a 2 (perpetrator: present, absent) X 2 (shirt: same, different) X 2 (hat: match, mismatch) between-subjects factorial design. The dependent variables were correct

identifications of the guilty suspect, false identifications of the innocent suspect, and confidence ratings.

Materials

Two mock crime videos were constructed for this experiment. The videos were highly similar except in one the perpetrator wore a hat and in the other he wore no hat. The perpetrator was a college age, clean-shaven, white male, with medium build and dark hair. The perpetrator wore a white t-shirt and blue jeans. The videos depicted the perpetrator entering a living room, looking down a hallway, and then stealing a laptop computer from a coffee table before quickly running out. The perpetrator's face was intermittently in view of the camera for approximately 10 s and the entire video lasted approximately 30 s.

The following eight showup identification conditions were created: same shirt-hat match, same shirt-hat mismatch, different shirt-hat match, and different shirt-hat mismatch, each for a perpetrator present and a perpetrator absent version. In the same shirt conditions, the guilty and innocent suspects wore the same jeans and white t-shirt worn by the perpetrator. In the different shirt conditions, the suspects wore the same jeans, but a gray hooded sweatshirt replaced the white t-shirt. The hat manipulation was very similar to the clothing manipulation. The hat match conditions include when the hat presence matched between video and identification (hat was worn in the video and identification and when the hat was neither worn in the video nor identification). The hat mismatch conditions include when the hat appeared in the video or identification but not both. Each showup identification was conducted as a video that lasted approximately 10 s. Each was filmed in a parking lot, with bushes and empty parking spaces in the background; the suspect stood still looking directly at the camera.

Procedure

All participants took part in this study online via Surveymonkey, which served to minimize experimenter bias and ensure random assignment. The informed consent made participants aware that this was an eyewitness identification study. After providing consent and demographic information, participants viewed the mock crime video. Following the video, participants completed 20 anagrams of insect names (e.g., REOPSSAHGRP). Participants were told to spend approximately 20 s on each word, for a total of 5 minutes on the distractor task. After the distractor task, participants were randomly assigned to one of the eight showup conditions. All participants were given unbiased instructions warning them that the perpetrator may or may not be in the showup. Following the instructions, participants made their identification decision by indicating "yes", that was the person from the video, or "no", that was not the person from the video. Participants then were asked to report how confident they were that the identification decision was correct on a scale anchored at 1 and 10, with 1 being "not at all confident" and 10 being "extremely confident". The following questions were then answered: Was the individual who committed the crime and the individual in the identification task wearing the exact same thing? What was the individual committing the crime wearing? What was the individual in the identification procedure wearing? Was the individual in the identification procedure wearing a hat? Was the individual committing the crime wearing a hat? Where in the home did the crime take place? These questions served as a manipulation check, and the last two were the designated exclusion questions. Participants were debriefed after completing these questions.

Results

The goal of this study was to compare showup identification accuracy as a function of the degree of match due to shirt and hat. To assess this, we calculated the correct and false

identification rates for each condition, along with the corresponding signal-detection response bias (criterion estimates, *c*). We conducted ROC analyses to ascertain if discrimination differed among the match and mismatch conditions of the shirt and hat. Finally, logistic regression analyses were conducted to determine the contributors to the discriminability and response bias differences.

Identification Rates

Table 1 displays the choosing rates by shirt and hat conditions (see Appendix A for frequencies of suspect identifications by confidence). The same shirt conditions, collapsed across match and mismatch hat conditions, yielded a correct identification rate of 48.8% and a false identification rate of 18.6%; the different shirt conditions, collapsed across match and mismatch hat conditions, yielded a correct identification rate of 27.4% and false identification rate of 16.5%. This corresponds to an estimated *c* of 0.47 for the same shirt conditions compared to a more conservative response bias of 0.79 for the different shirt conditions.

The hat match conditions, collapsed across same and different shirt, yielded a correct identification rate of 44.9% and a false identification rate of 21.3%; the hat mismatch conditions, collapsed across same and different shirt, yielded a correct identification rate of 29.3% and false identification rate of 13.5%. This corresponds to an estimated *c* of .46 for the hat match conditions compared to a more conservative response bias of .82 for the hat mismatch conditions.

Further examination of Table 1 reveals that the highest correct and false identification rates occurred when the cues at encoding (i.e., the video) matched those presented at retrieval (i.e., showup). Because we varied both the shirt and hat, we were able to produce pure conditions of match (the shirt and hat match at encoding and test) and neither match (neither the shirt nor

hat match between encoding and test). In addition, we were able to compare two partial match conditions in which either the shirt or the hat was the matching feature. A complete match, the same shirt and hat match at showup, yielded higher correct identification rates (58.9%) than did partial matches of same shirt-hat mismatch and different shirt-hat match (36.7% and 32.0%, respectively). The neither match condition (different shirt and hat mismatch) yielded the lowest correct identification rate at 22.7%. When we consider response bias, we see the most liberal responding in the complete match condition (c = 0.06), which was more liberal than the different shirt-hat match condition (c = 0.67), followed by the same shirt-hat mismatch condition (c = 0.72), with the neither match condition being the most conservative (c = 0.93).

This initial evaluation of response bias based on the separate correct and false identification rates indicates that a clothing match, more specifically the same shirt and hat match, affects response bias in that participants responded more liberally. There was little difference in response bias among the two partial match and neither match conditions, all of which were more conservative than the complete match condition.

ROC Analysis

Although we can evaluate response bias by assessing correct and false identification rates, it can be misleading to use these values to assess discriminability (Gronlund & Neuschatz, 2014; Mickes, Moreland, Clark, & Wixted, 2014). Although a measure like d' can be used, Mickes et al. (2014) point out that it requires assumptions that likely are violated here. Therefore, to evaluate discriminability differences, we rely upon Receiver Operating Characteristic (ROC) analysis, which is assumption-free. ROC analysis traces out discriminability at different levels of response bias, or confidence (Gronlund, Wixted, & Mickes, 2014; Wixted & Mickes, 2012). The point defined by the correct and false identification rates in the lower left portion of the ROC

curve reflect those suspect identifications made with the highest level of confidence. The next point on the ROC reflects the correct and false identifications made at the two highest confidence levels. This continues to the right-most point on the ROC, which reflects the overall correct and false identification rates made at all levels of confidence. The condition whose ROC curve lies closest to the upper left-hand corner of the ROC space (farthest from the chance diagonal) is the condition that produces the best discriminability between the guilty and innocent suspect. Discriminability for the various conditions is evaluated statistically by computing the area under the curve (AUC) for each condition (see Gronlund et al., 2014).

We begin by comparing showup performance for same versus different shirt, collapsing over whether the hat was a match or mismatch at identification. We then proceed to comparing performance for the hat, collapsing over whether the shirt was the same or different, and lastly to the conditions of complete match, both partial match, and neither match, between the video and showup.

Discriminability was significantly better in the same shirt (AUC = .65) than the different shirt condition (AUC = .56), $D = 5.61$, $p < .001$. The positive effect of a shirt match is contrary to the generally accepted conclusion that a clothing match has a negative effect on discriminability (Yarmey et al., 1996; Dysart et al., 2006). However, there was no significant difference in the AUCs for the hat match (.63) versus hat mismatch (.59), $D = 1.32$, $p = .18$, although the curves fall in the predicted direction.

The top panel of Figure 1 displays the ROC curves for the complete match, two partial match, and neither match comparisons. The complete match AUC (.69) was significantly greater than the same shirt-hat mismatch (.61), different shirt-hat match (.57), and neither match conditions (.55), AUCs, $D = 2.85$, $p < .004$, $D = 4.89$, $p < .001$, and $D = 5.93$, $p < .001$,

respectively. The same shirt-hat mismatch AUC was significantly greater than either the different shirt-hat match or neither match AUCs, $D = 2.17$, $p < .05$ and $D = 3.15$, $p < .001$, respectively. These results indicate that the best discriminability results when the cues available during the video (encoding) match the cues provided during the showup (retrieval).

Logistic Regression

Now that we have established that discriminability differences exist among clothing match conditions, we can determine whether correct or false identifications (or both) are driving the discriminability differences. To accomplish this we conducted logistic regression analyses on correct and false identifications separately.

Correct Identifications. Logistic regression analysis was used to examine correct identifications using shirt (same, different), and hat (match, mismatch) as predictors. All *p*-values were two-tailed unless otherwise noted. These variables significantly predicted correct identifications. The two-way interaction was significant, *Wald* X^2 (1, $N = 584$) = 23.45, $p <$ .001.There were also significant main effects for shirt (*Wald* X^2 (1, $N =584$) = 26.34, $p < .001$ and hat, *Wald* X^2 (1, $N =584$) = 9.35, $p = .002$.

A priori chi-squares were conducted to explore the predictions made by the ICE theory. We utilized the same four match conditions discussed previously: complete match, same shirt-hat mismatch, different shirt-hat match, and neither match. We compared the complete match condition to all other conditions, and found that there were significantly more correct identifications in the complete match condition than any other conditions (all X^2*s* (1, $N = 306$) > 4.39, $p < .036$, $\Phi > .124$). In addition, in the neither match condition, participants had significantly fewer correct identifications than in any other condition, all X^2*s* (1, $N = 299$) > 3.89, $p < .049$, $\Phi > -.114$.

False Identifications. We next examined the false identification rates using the logistic regression comparisons described above. Consistent with the correct identifications, there was a significant interaction of shirt and hat, *Wald* X^2 (1, N = 601) = 5.48, p = .019. There was a main effect of hat, *Wald* X^2 (1, N = 601) = 4.02, p = .045, but not shirt, *Wald* X^2 (1, N = 601) = 1.004, p = .316. There was a significant difference between the complete match and neither match conditions, indicating that there were more false identifications in the complete match than the neither match conditions (X^2 (1, N = 284) = 4.52, p = .033, Φ = .126). There were also significantly more false identifications in the complete match than the same shirt-hat mismatch, X^2 (1, N = 277) = 4.38, p = .036, Φ = .126. There was no difference between the complete match and the different shirt-hat match, X^2 (1, N = 318) = 2.062, p = .151.

Discussion

We examined the effect of clothing match on showup identification performance. Participants were better able to discriminate the guilty suspect from the innocent suspect when the suspects wore the same shirt in the video and the showup than when wearing a different shirt. Moreover, the complete match conditions yielded a higher ROC curve than either partial match or neither match comparisons. Overall, the data are consistent with the ICE theory (Murnane et al., 1999). Participants likely created an ensemble, or compound cue, involving the face, clothing, and hat, and were best able to make an accurate identification when the ensemble matched. The logistic regression analyses revealed a liberal response bias shift in the complete match condition, but those who were choosing were still better able to discriminate. Although it appears that participants were able to form an ensemble, the associated context also influenced their identification decisions. These results are consistent with previous research that indicated an increase in false identifications from clothing match (Yarmey et al., 1996; Dysart et al., 2006),

which the ICE theory would consider due to the contribution of associated context. The current research demonstrates the importance of examining both the false identifications and correct identifications to differentiate discriminability from response bias differences. Although on the surface our results may appear contrary to the generally accepted conclusion that a clothing match harms discriminability, there are reasons to believe that our conclusions are reliable. First, we examined how shirt and hat affect both the correct and false identification rates; past research largely focused on examining false identification rates. Secondly, we used ROC analyses to assess the discriminability of the different clothing conditions. Reliance on correct and false identification rates, or ratios of these rates, can be misleading because they conflate discriminability and response bias (Wixted & Mickes, 2012). Additionally, we utilized a much greater sample size than prior research, providing sufficient power (Maxwell, Kelley, & Rausch, 2008). Lastly, we did not encounter ceiling effects, which may have hampered past research (Dysart et al., 2006; Smith et al., 2012).

Limitations

One possible limitation of the present study is that the perpetrator and innocent suspect wore common clothing, and a possible negative effect of clothing match might result from the use of distinctive or atypical clothing. However, the data on this point are mixed. Dysart et al. (2006) found an increase in false identifications when distinctive clothing was worn during the video and subsequent showup identification, but found no effect on correct identifications, and were unable to find an effect of clothing match with common clothing. However, Valentine et al. (2012) failed to find an increase in false identifications when distinctive (or what the authors referred to as “loud”) clothing was worn. Thus, it is unclear if the effects of clothing on false identifications would have been any larger if distinctive clothing were used in the present study.

Another possible limitation is that only one guilty suspect was employed in the current study. This raises the concern that the results may be idiosyncratic to the particular suspect and lacking in external validity (Wells & Windschitl, 1999). To address this concern, we collected data for an additional 271 participants in a hat video condition in which the roles of the innocent suspect and the perpetrator were switched. The video and the materials were exactly the same other than the switched roles of the guilty and innocent suspect. The resulting correct and false identification rates are italicized in Table 1. Individual 2 x 2 x 2 hierarchical loglinear analyses (HILOG) were performed to examine the influence of suspect (suspect A, suspect B), shirt (same, different) and hat (match, mismatch) on participants' identification accuracy (correct, incorrect) for the perpetrator present and perpetrator absent conditions. There was no effect of suspect in the perpetrator present (χ^2 (1) = .013, p = .91) or perpetrator absent (χ^2(1) = .349, p = .55) conditions, nor did suspect interact with any variables. In fact, inspection of Table 1 shows that the pattern of results is unchanged for either suspect pair; the highest correct and false identifications occur when there is the greatest match between encoding (crime video) and retrieval (showup).

Lastly, the contrived nature of the showup identification task is of note. As stated previously, in real-life showups, suspects are caught near the scene of a crime with respect to both distance and time, and presented live to an eyewitness. The social interaction between the witness/victim and the showup administrator are also important factors to consider. The video showups used in the current study did not try to simulate these influences, although they did provide additional retrieval cues beyond what static photographs provide.

Practical Applications

Maintaining the same clothing in the showup identification enhanced participant's ability

to discriminate a guilty from an innocent suspect. Does that mean that we recommend that the police should try to capitalize on a match of clothing and not worry about any possible negative consequences? An answer to that question depends on the quality of an eyewitness' description of the perpetrator. For example, if the viewing conditions are suitable (e.g., adequate lighting, low stress) and a witness is able to provide a detailed description of the perpetrator's face, then it seems reasonable that a clothing match could enhance discriminability. The witness would be more likely to create an ensemble and consequently the match would increase discriminability. However, if the eyewitness reports getting a poor look at a perpetrator, resulting in poor memoryfor the face, an item of clothing could be the primary feature encoded. For instance, if the perpetrator's face is in view for a few seconds, is seen from a distance, or is wearing a disguise (e.g., a hat covering part of his face), there may not be enough time to encode an ensemble that contains the face and the clothing. Furthermore, a stressful situation like a crime may make the formation of an ensemble difficult due to the cognitive effort required, making clothing more likely to function as associated context, thereby inducing only a liberal response bias shift without the increased discriminability. Before a strong recommendation regarding the appropriateness of relying on a clothing match at showup identifications can be made, further research needs to be conducted to thoroughly examine the circumstances under which a match would provide a memorial advantage.

References

Cutler, B. L., Penrod, S. D., & Martens, T.K. (1987). Improving the reliability of eyewitness identification: Putting context into context. *Journal of Applied Psychology, 72*, 629-637. doi:10.1037/0021-9010.72.4.629

Cutler, B.L., Penrod, S.D., O'Rourke, T.E., & Martens, T.K. (1986). Unconfounding the effects of contextual cues on eyewitness identification accuracy. *Social Behaviour, 1*, 113-134.

Dysart, J. E., & Lindsay, R. C. L. (2007). Show-up identifications: Suggestive technique or reliable method? In Rod C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *The handbook of eyewitness psychology, Vol II: Memory for people.* (pp. 137-153). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Dysart, J. E., Lindsay, R. C. L., & Dupuis, P. R. (2006). Show-ups: The critical issue of clothing bias. *Applied Cognitive Psychology, 20,* 1009-1023. doi:10.1002/acp.1241

Goodsell, C. A., Wetmore, S. A., Neuschatz, J. S., & Gronlund, S. D. (2013). Showups. In B. L. Cutler (Ed.), *Reform of eyewitness identification procedures.* (pp. 45–63). Washington, DC US: American Psychological Association.

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

Gronlund, S. D., Goodsell, C. A., & Andersen, S. M. (2012). Lineup procedures in eyewitness identification. L. Nadel & W. Sinnott-Armstrong (Eds.), *Memory and Law*, New York: Oxford University Press.

Gronlund, S. D., Wixted, J. T., & Mickes, L. (2014). Evaluating eyewitness identification procedures using ROC analysis. *Current Directions in Psychological Science, 23,* 3-10. doi:10.1177/096372143498891

Gronlund, S. D., & Neuschatz, J. S., (2014). Eyewitness Identification discriminability: ROC analysis versus logisitic regression. *Journal of Applied Research in Memory and Cognition, 3,* 54-57. doi:10.1016/j.jarmac.2014.04.008

Lawson, V. Z., and Dysart, J. E. (2012). The showup identification procedure: An exploration of systematic biases. *Legal and Criminological Psychology*. doi:10.1111/j.2044-8333.2012.02057.x

Lindsay, R. C., Wallbridge, H., & Drennan, D. (1987). Do the clothes make the man? An exploration of the effect of lineup attire on eyewitness identification accuracy. *Canadian Journal of Behavioural Science, 19,* 463-478. doi:10.1037/h0079998

Maxwell, S. E., Kelley, K., & Rausch, J. R. (2008). Sample size planning for statistical power and accuracy in parameter estimation. *Annual Review of Psychology, 59,* 537-563. doi:10.1146/annurev.psych.59.103006.093735

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

Mickes, L., Moreland, M. B., Clark, S. E. & Wixted, J. T. (2014). Missing the information needed to perform ROC analysis? Then compute d' not the diagnosticity ratio. *Journal of Applied Research in Memory and Cognition, 3,* 58-62. doi:10.1016/j.jarmac.2014.04.007

Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context-dependent recognition memory: The ICE theory. *Journal of Experimental Psychology: General, 128,* 403-415. doi:10.1037/0096-3445.128.4.403

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory, 10,* 389-395. doi:10.1080/09658210244000216

Neuschatz, J. S., Wetmore, S. A., Key, K. N., Cash, D. K., Gronlund, S. D., & Goodsell, C. A. (in press). Comprehensive evaluation of showups. In M. Miller & B. Bornstein (Eds.), *Advances in Psychology and Law*. New York: Springer.

Schmechel, R. S., O'Toole, T. P., Easterly, C., & Loftus, E. F. (2006). Beyond the ken? Testing jurors' understanding of eyewitness reliability evidence. *Jurimetrics, 46,* 177-214.

Smith, S. M. (1988). Environmental context-dependent memory. In Davies, G. M., & Thomson, D. M. (Eds.), *Memory in context: Context in memory* (pp. 231-244). New York, NY: John Wiley & Sons.

Smith, S. M. (1994). Theoretical principles of context-dependent memory. In Morris, P., & Gruneberg, M. (Eds.), *Theoretical aspects of memory* 2nd Edn. (pp. 168-195). New York, NY: Routledge Press.

Smith, A. M., Leach, A. M., & Cutler, B. L. (2013). Facilitating accuracy in showup identification procedures: The effects of the presence of stolen property. *Applied Cognitive Psychology, 27,* 216-221. doi:10.1002/acp.2898

Sporer, S. L. (1993). Clothing as a contextual cue in facial recognition. *German Journal of Psychology, 17,* 183-199.

Stovall v. Denno, (1967). 388 U.S. 293.

Susa, K. J. & Meissner, C. A. (2008). Accuracy of eyewitness descriptions. In B. Cutler (Ed.), *Encyclopedia of Psychology & Law, Vol. 1* (pp. 285-287). Sage Publications.

Technical Working Group for Eyewitness Evidence. (1999). Eyewitness evidence: A guide for law enforcement. Washington, D.C: United States Department of Justice, Office of Justice Programs.

Tulving, E., & Thomson D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352-373. doi:10.1037/h00020071

Valentine, T., Davis, J. P., Memon, A., & Roberts, A. (2012). Live showups and their influence on a subsequent video line-up. *Applied Cognitive Psychology, 26,* 1-23. doi:10.1002/acp.1796

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 14,* 442-452. doi: 10.1037/0278-7393.1.4.442

Watkins, M. J, & Watkins, O. C. (1976). Cue-overload theory and the method of interpolated attributes. *Bulletin of the Psychonomic Society, 7,* 289-291. doi:10.3758/BF03337192

Wells, G. L. & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin, 25* (9), 1115-1125.

Wetmore, S. A., Neuschatz, J. S., Gronlund, S. D., Goodsell, C, A., Wooten, A., & Carlson, C. A. (in press). Are showups ever better than lineups? *Journal of Applied Research in Memory and Cognition.* doi:10.1016/j.jarmac.2014.07.003

Wixted, J. T., & Mickes, L. (2012). The field of eyewitness memory should abandon probative value and embrace receiver operating characteristic analysis. *Perspectives in Psychological Science, 7,* 275-278. doi:10.1177/1745691612442906

Yarmey, A. D., Yarmey, M. J., & Yarmey, A. L. (1996). Accuracy of eyewitness identifications in showups and lineups. *Law and Human Behavior, 20,* 459-477. doi:10.1007/BF01498981

Table 1

Response Percentages by Same or Different Shirt and Hat Match or Mismatch

	Same Shirt Hat Match (Complete)	**Same Shirt Hat Mismatch (Partial)**	**Different Shirt Hat Match (Partial)**	**Different Shirt Hat Mismatch (Neither)**
Perpetrator Present				
ID	58.9 (182)	36.7 (109)	32.0 (107)	22.8 (76)
N	309	297	334	334
ID*	*75.0 (24)*	*43.8 (14)*	*42.1 (16)*	*30.8 (8)*
*N**	*32*	*32*	*38*	*26*
Perpetrator Absent				
ID	24.2 (73)	13.1 (40)	18.8 (65)	13.8 (119)
N	302	305	345	312
ID*	*36.4 (12)*	*27.9 (12)*	*30.3 (10)*	*8.8 (3)*
*N**	*155*	*166*	*187*	*182*

Note: Frequency for each condition is presented in parentheses. ID = identification of suspect. * = Reversal of perpetrator and innocent suspect roles.

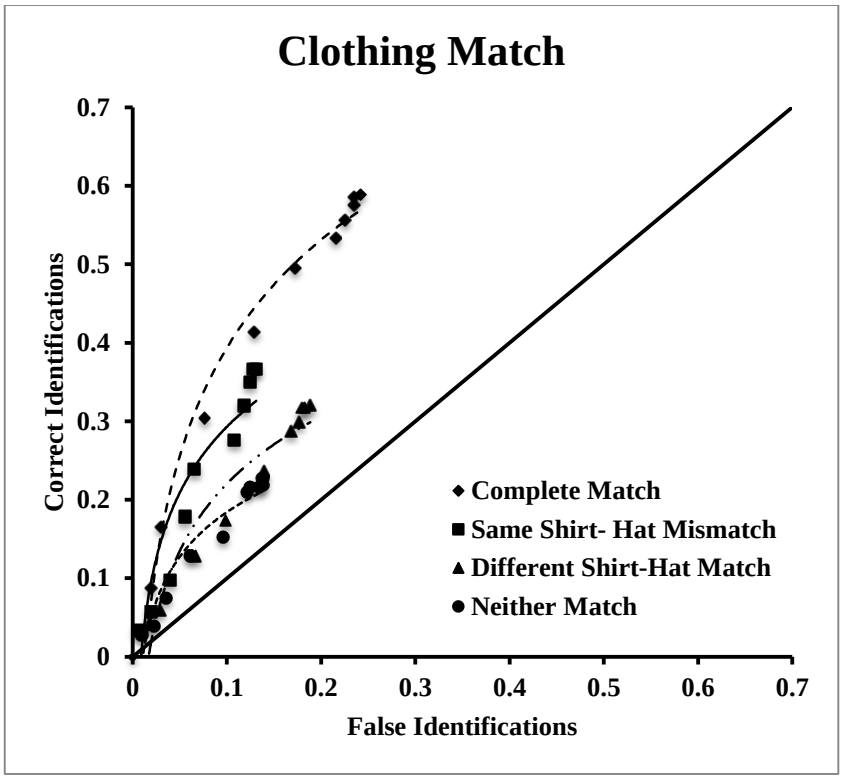

Figure 1. ROC curves comparing the complete match, same shirt-hat mismatch, different shirt-hat match, and neither match conditions between video and showup. The diagonal line indicates chance performance.

Appendix A

Frequencies of Suspect Identifications as a Function of Confidence Level

	Same Shirt Hat Match (Complete)		**Same Shirt Hat Mismatch (Partial)**		**Different Shirt Hat Match (Partial)**		**Different Shirt Hat Mismatch (Neither)**	
	PP	**PA**	**PP**	**PA**	**PP**	**PA**	**PP**	**PA**
Confidence								
10	27	6	10	2	14	6	9	3
9	24	3	7	4	6	4	4	4
8	43	14	12	6	23	13	12	4
7	34	16	24	5	15	11	18	8
6	25	13	18	3	21	14	8	11
5	12	13	11	13	17	10	19	8
4	7	3	13	3	4	3	2	1
3	6	3	9	2	6	1	0	3
2	3	0	5	1	0	1	1	1
1	1	2	0	1	1	2	3	0

Note. PP is perpetrator present and PA is perpetrator absent.