Accepted for publication in *Law and Human Behavior*

November 2005

Please do not quote until final technical editing is completed

Eyewitness Lineups: Is the Appearance-Change Instruction a Good Idea?

Steve D. Charman and Gary L. Wells

Iowa State University

Corresponding author:

Steve D. Charman

Psychology Department

West 112 Lago

Iowa State University

Ames, IA 50011

e-mail: scharman@iastate.edu

Abstract

The Department of Justice's Guide for lineups recommends warning eyewitnesses that the culprit's appearance might have changed since the time of the crime. This appearance-change instruction (ACI) has never been empirically tested. A video crime with four culprits was viewed by 289 participants who then attempted to identify the culprits from four 6-person arrays that either included or did not include the culprit. Participants either received the ACI or not and all were warned that the culprit might or might not be in the arrays. The culprits varied in how much their appearance changed from the video to their lineup arrays, but the ACI did not improve identification decisions for any of the lineups. Collapsed over the four culprits, the ACI increased false alarms and filler identifications but did not increase culprit identifications. The ACI reduced confidence and increased response latency. Two processes that could account for these results are discussed, namely a decision criterion shift and a general increase in ecphoric similarity.

Eyewitness Lineups: Is the Appearance-Change Instruction a Good Idea?

Consider two witnesses to a crime who had equally good views of the culprit. Each agrees to view a photo lineup and each is warned that the actual culprit might not be in the lineup. The first witness views the lineup and does not see anyone who fits her memory of the culprit well enough to make a positive identification. For the second witness, the detective administering the lineup adds a phrase to the instruction: “Keep in mind that the culprit’s appearance may have changed and they might not appear exactly as they did at the time of the crime.” The second witness has the same initial reaction as the first witness but, recalling the instruction, realizes that an exact match in appearance is not needed. The second witness reasons that if number three’s hair were cut and combed more like the culprit’s hair and some weight taken off of his face, then he would look a lot like the culprit. The second witness decides to identify number three from the lineup.

Assume now that the only difference between the two witnesses was the appearance-change instruction given to the second witness. Did this appearance-change instruction help the identification process? If the actual culprit was in the lineup, and was in fact number three, then the instruction served a beneficial function by causing the witness to identify the culprit. If the actual culprit was not in the lineup, however, the instruction served a detrimental function by causing a mistaken identification.

The use of such an appearance-change instruction was part of a set of guidelines prepared by a Department of Justice working group on the collection of eyewitness evidence and given to all law enforcement agencies in the United States in 1999 (Technical Working Group, 1999, hereafter called the Guide in this article). Specifically, the Guide recommends that eyewitnesses be instructed before viewing a lineup “that individuals depicted in lineup photos may not appear exactly as they did on the date of the incident because features such as head and facial hair are subject to change” (Technical Working Group, 1999, p. 32). Although most of the Guide’s recommendations were based on eyewitness research, the appearance-change instruction is an exception; no prior research has tested the properties of the appearance-change instruction.

The absence of research on the appearance-change instruction stands in marked contrast to considerable research on the other main pre-lineup instruction recommended in the Guide, namely the instruction that the person who committed the crime might or might not be in the lineup. First tested by Malpass and Devine (1981), the might-or-might-not-be-present instruction (hereafter called the present/absent instruction) has been subjected to a recent meta-analysis that indicates that the present/absent instruction reduces mistaken identifications when the culprit is absent but also reduces correct identifications when the culprit is present (Clark, 2005). Nevertheless, the present/absent instruction reduces mistaken identifications from culprit-absent lineups at a considerably greater rate than it reduces correct identifications from culprit-present lineups. Hence, there is a general consensus among eyewitness researchers that the present/absent instruction is an essential component of good pre-lineup instructions. This consensus is evidenced by surveys of eyewitness experts (e.g., see Kassin, Tubb, Hosch, & Memon, 2001) and the explicit incorporation of the present/absent instruction in the recommendations of the American Psychology-Law Society review of lineups (Wells et al., 1998).

Conceptually, the present/absent instruction is very different from the appearance-change instruction. The present/absent instruction fits nicely within the theoretical and structural frameworks that have guided the lineup literature (Charman & Wells, in press). Specifically, because of the single-suspect structure of lineups, mistaken identifications of innocent suspects (rather than known-innocent fillers) occur when the culprit is not present (Wells & Turtle, 1986)[1]. When eyewitnesses are forced by the present/absent instruction to consider the possibility that the person who looks most like the culprit is not the culprit, the tendencies of eyewitnesses to rely on mere relative judgments (selecting the person who looks most like the culprit) are presumed to be diminished (Wells, 1984). Correct rejections, therefore, depend heavily on making sure that the eyewitness considers the possibility that the culprit is not in the lineup. Hence, both the structure of lineups and the psychological processes that lead to mistaken identifications are highly congruent with the present/absent instruction. The appearance-change instruction, on the other hand, does nothing in itself to sensitize the eyewitness to the possibility that the culprit might not be present. Nor is there any theoretical reason to speculate that the appearance-change instruction would reduce reliance on relative judgments.

The absence of data on the appearance-change instruction is thus matched by the absence of theory concerning the instruction. Nevertheless, we can extract what we believe is the implicit theory behind the appearance-change instruction. The implicit theory seems to be that correct identifications from culprit-present lineups are lost, at least in part, because the culprit's appearance in the lineup is not exactly as the culprit appeared during witnessing whereas the eyewitness expects a more perfect match. To the extent that the culprit's appearance has in fact changed, the appearance-change instruction might permit the witness to see beyond the change and recognize the culprit. Hence, the implicit theory is that the appearance-change instruction will help to reduce misses (i.e., incorrect rejections or failures to identify the culprit from culprit-present lineups). This is a plausible theory and, if true, we would expect this effect to be moderated by actual appearance change: The greater the appearance change, the more beneficial the appearance-change instruction should be for reducing misses when the culprit is present.

There has been little work in the eyewitness literature directly concerning the effects of a criminal's naturalistic appearance change on identification rates. There are, however, numerous studies on the effects of disguise and a meta-analysis has shown a strong debilitating effect ($d = 1.05$) of disguise on accurate recognition rates (Shapiro & Penrod, 1986). Disguise, however, is not quite the same as naturalistic appearance change. Studies of disguise typically involve the addition and deletion of external factors, such as hats, eyeglasses, or masks. Naturalistic appearance change, in contrast, derives from changes in appearance that occur with age, hair style, weight, and other appearance factors that occur to greater or lesser extents in people because of the passage of time. Nevertheless, the strong effects of disguise suggest that dissimilarities between the appearance of the culprit at the time of witnessing and at the time of the lineup will have strong debilitating effects on hit rates, whether they come from naturalistic change or from disguise. Thus, the appearance-change instruction might help prevent some of this loss in hit rates. But, are there any additional problems that might be introduced with the appearance-change instruction?

It has long been recognized that the net value of any lineup manipulation must be tested under conditions in which the culprit is absent as well as conditions in which the culprit is present in the lineup (Lindsay & Wells, 1980). Some models of the lineup identification process clearly predict that lineup manipulations that increase hit rates (reduce misses) for culprit-present lineups will generally also increase mistaken identifications (reduce correct rejections) for culprit-absent lineups. The magnitudes of the increase in hit rate versus increase in mistaken identification rate need not be the same, but some trade-off is presumed.

Take, for example, Clark's WITNESS model, a mathematical model of eyewitness identification performance from lineups, which has been shown to fit experimental data quite well (Clark, 2003). The WITNESS model assumes that a lineup identification will occur when the weighted sum of two different judgments surpasses the witness's identification criterion. These two judgments, based on Wells' (1984) distinction between absolute and relative judgments, are in turn wholly determined by the various matches between the appearance of each lineup member and the eyewitness's memory of the culprit. The "BEST" judgment, analogous to absolute judgments, represents the match value of the one lineup member that looks most like the witness's memory of the culprit. The "DIFF" judgment, analogous to relative judgments, represents the difference in match values between the best and next-best lineup members. We will refer to the match between the eyewitness's memory of the culprit and the appearance of a lineup member as "ecphoric similarity," a term used by Tulving (1981) that refers to the extent to which a stimulus resembles a person's memory of that stimulus. Under Clark's framework, the probability of an identification is purely a function of the relation between the various ecphoric similarities of the lineup members and the witness's decision criterion. Hence, lineup manipulations that affect eyewitness accuracy are the result of changing one of these two variables. For example, clothing bias, in which the suspect-at-lineup is dressed like the culprit, has been shown to increase identifications of the suspect, presumably because of an increase in ecphoric similarity. Sequential lineups (in which lineup members are shown one at a time to the witness) have been shown to decrease both false identifications and correct identifications when compared to simultaneous lineups (in which lineup members are shown simultaneously to the witness); some researchers believe that this is a result of the sequential procedure raising a witness's decision criterion (Ebbeson & Flowe, 2002). Recent research using the Signal Detection Theory framework is also consistent with the view that the sequential procedure produces a more conservative decision criterion (Meissner, Tredoux, Parker, & MacLin, 2005).

The WITNESS model itself is silent regarding the causal variables and processes that raise and lower the decision criterion and ecphoric similarity. Hence, we can only speculate at this point about how the appearance-change instruction might impact the two WITNESS model factors. First, we consider the possible effect of the appearance-change instruction on the decision criterion. The presumption of the WITNESS model is that the decision to not make an identification results from situations in which no lineup member has enough ecphoric similarity to surpass the decision criterion. But the decision criterion itself is presumed to depend on the amount of ecphoric similarity the eyewitness would expect to see if the culprit is in the lineup. An appearance-change instruction could very well lower the decision criterion because it implies to the eyewitness that the culprit might not evoke much ecphoric similarity due to an appearance change. If the appearance-change instruction lowers eyewitnesses' expectations regarding how much ecphoric similarity to expect, eyewitnesses might lower their decision criterion. If so, we would expect an appearance-change instruction to increase rates of choosing for both culprit-present as well as culprit-absent lineups.

Could the appearance-change instruction also affect ecphoric similarity? This is a more complex issue. Ecphoric similarity is the extent to which a stimulus resembles one's memory for the stimulus. In a lineup task, this is usually thought of as the extent to which a lineup member's appearance in the lineup resembles the eyewitness's memory of the culprit. Clearly, the appearance-change instruction cannot affect the lineup member's actual appearance. And we do not think it plausible to suggest that the appearance-change instruction affects the eyewitness's memory of the culprit. So it seems odd to suggest that the appearance-change instruction could affect ecphoric similarity. However, one of the possible consequences of the appearance-change instruction is to lead witnesses to engage in various mental mutations to the appearance of the various lineup members. Our speculation about mental mutations is derived in part from the simulation heuristic, originally put forth by Kahneman and Tversky (1982), which describes how people can take a factual event and imagine alternative values for different aspects of the event. Hence, eyewitnesses could mentally alter the hair, add a beard, or imagine the culprit at a heavier weight. The mental image resulting from this mutation could then become a new (internal) stimulus that is used to make judgments of ecphoric similarity. Because the number of these mutated mental images is limited only by the imagination of the witness, it is highly likely that at least one of them will resemble the witness's memory of the culprit more than the actual photo resembles the witness's memory of the culprit. Thus, by encouraging these mental mutations, the appearance-change instruction can, in effect, increase ecphoric similarity by changing the stimulus to which one's memory of the culprit is compared.

Importantly, however, this begs the question of whether or not the net result of the appearance-change instruction is an increase in accurate identification rates without an increase in mistaken identifications. Presumably, these mutations would be applied to all lineup members with the net result perhaps being an increase in ecphoric similarity for every lineup member. If so, then we might expect an increase in the overall rate of identifications (a reduction of "not there" responses) but not necessarily in a manner that selectively elevates identifications of the perpetrator. In fact, for culprit-absent lineups, increases in ecphoric similarity are not desirable because such increases are likely to lead to increased chances of misidentification. Hence, the issue is whether increased ecphoric similarity is selective to the culprit or whether the increase also applies to innocent lineup members. We will call the former process a "selective increase in ecphoric similarity" and the latter a "general increase in ecphoric similarity."

We think that it is plausible that the appearance-change instruction could produce a selective increase in ecphoric similarity or, at least, a relatively greater increase in apparent ecphoric similarity for the culprit than for innocent lineup members. It is, after all, the culprit's face that the witness originally encoded and it is the culprit's face that has undergone some change in appearance. Therefore, these mental mutations could make the culprit closer to the witness's memory than would these same mutations to innocent lineup members, resulting in a selective increase in ecphoric similarity. Despite the plausibility of this account, there are two main reasons why we did not predict a selective increase in ecphoric similarity from the appearance-change instruction. First, it seems to us that a selective increase in ecphoric similarity depends in part on the witness knowing what kind of appearance change the culprit has undergone, if any. Furthermore, we view the mental mutation process as a form of hypothesis testing and we know that people generally test hypotheses in a confirmation-biased manner (see Nickerson, 1998). Hence, we would expect witnesses in culprit-absent conditions to routinely mutate hair, age, weight and other changeable characteristics of the lineup members in the direction of their memory for the culprit. This, in turn, should lead to increases in ecphoric similarity for the innocent lineup members as well as the culprit, resulting in a general increase in ecphoric similarity.

Because there is no prior research on the appearance-change instruction with lineups, the primary goal of this research was restricted to finding out whether the net effect of the instruction was to increase identifications of the culprit without also increasing the rate of identifications for innocent lineup members. Although we have described three theoretical processes that could affect the results (a selective increase in ecphoric similarity, a general increase in ecphoric similarity, and a lowering of response criterion), we cannot necessarily distinguish between these process explanations, depending on the results. For example, if choosing rates increase for both the culprit-present and culprit-absent conditions with no net increase in accuracy, then this could be due to either a general increase in ecphoric similarity or a lowered response criterion. In that case, we will have ruled out only one process, namely a selective increase in ecphoric similarity. If, on the other hand, the increase in choosing rates is uniquely elevated for the culprit and not for the innocent lineup members, then we can rule out a general increase in ecphoric similarity and a lowered response criterion. Also note that although a selective increase in ecphoric similarity and a general increase in ecphoric similarity are mutually-exclusive processes, either one of these processes can occur in conjunction with a lowered decision criterion. Thus, evidence for a change in ecphoric similarity should not necessarily be interpreted as evidence against a lowered decision criterion (and vice versa).

Any test of the appearance-change instruction clearly requires that there be culprits that vary in appearance change. Accordingly, we used a video crime in which there were four culprits whose lineup photos resembled their appearance in the video to significantly different degrees. This allowed us to test the effects of the appearance-change instruction using culprits who had undergone very little appearance change, as well as culprits who had undergone considerable appearance change. We did not manipulate appearance change within culprits. At one level, this limits our conclusions because we cannot say that differences in hit rates across targets are due solely to appearance change. However, the focus of this research is not on the effects of appearance change itself. Instead, the focus is on whether the appearance-change instruction improves hit rates without elevating false alarm rates. Using multiple culprits, including some that had undergone considerable amounts of appearance change, permitted circumstances for the appearance-change instruction to have a good chance of showing its effectiveness. Hence, it is more accurate to describe variations in appearance change across culprits as a form of strategic stimulus sampling than to describe it as a manipulation (see Wells & Windschitl, 1999).

In addition to the identification data, we collected two additional measures, namely confidence and response latency. We expected the appearance-change instruction to increase response latency regardless of which process (a change in ecphoric similarity or a change in decision criterion) the instruction evokes. For instance, if the appearance-change instruction leads witnesses to engage in mental mutations to the appearance of the lineup members, this additional process should add time to the witness's decision. However, even if the appearance-change instruction only lowers the decision criterion, the decision should take longer because it would increase the number of lineup members who fall above the criterion, resulting in a need to consider more lineup members to be viable than if the decision criterion were higher.

If the appearance-change instruction results in a selective increase in ecphoric similarity, then we expected that confidence in identification decisions would also increase. This is because a selective increase in ecphoric similarity should further separate the culprit from the fillers, which should make the identification decision easier. It was less clear what to expect with respect to confidence if either a general increase in ecphoric similarity or a lowering of decision criterion occurs. On the one hand, confidence might increase with the appearance-change instruction because discrepancies between the appearance of the identified person and the witness's memory could be dismissed by the notion of appearance change. On the other hand, if the appearance-change instruction leads to more lineup members becoming viable (either because more lineup members fall above the decision criterion or because of increased ecphoric similarity of multiple lineup members), the identification of any one of them might be made with less confidence. Notice that although confidence data may be used as evidence to support or refute the existence of a selective increase in ecphoric similarity, neither the response latency nor the confidence measures permit us to distinguish between a general increase in ecphoric similarity and a lowered decision criterion. Nevertheless, any changes in response latency, confidence, or both as a function of the appearance-change instruction helps to complete our description of the effects of the appearance-change instruction.

The appearance-change instruction is potentially a very important type of system-variable improvement for lineups. Other system-variable improvements for lineups (e.g., the present/absent instruction, the sequential procedure, selecting fillers who fit the description of the culprit) clearly are directed at reducing the chances of misidentification from culprit-absent lineups. No system-variable intervention has yet shown that it can reliably increase the rate of accurate identifications (reduce miss rates) from culprit-present lineups. This is important because both research experiments and archival analyses of real-world lineups indicate that "hit rates" typically run around 50% or less (for archival analyses, see Behrman & Davey, 2001; Valentine, Pickering, & Darling, 2003; Wright & McDaid, 1996; for experimental analyses see Cutler & Penrod, 1995). Hence, there is a great importance in determining the effects and potential usefulness of the appearance-change instruction.

Method

Pre-testing

One of the assumptions of the current study is that the effect of the appearance-change instruction will depend on the degree of appearance change that is undergone by a culprit between the time of the crime and the time of the lineup. Thus, to give the appearance-change instruction the best chance at producing beneficial effects, it was necessary to ensure that the four culprits used in the current study differed in the degree to which they experienced appearance change. In order to determine the degree of appearance change for each of the individual culprits, sixty-three university undergraduate participants provided similarity ratings during pre-testing. Pictures of six fillers were chosen for each of four culprits that appeared in the mock crime. Pre-testing participants viewed, in turn, the lineup photo of each member of the four lineups ([six fillers + one culprit] X four lineups = 28 lineup members) next to a looped five-second clip of the culprit to whom the lineup member corresponded. The five-second clip was taken directly from the mock crime video, and was pre-selected based on quality of view of the culprit's face. While viewing each lineup member next to the appropriate looped video clip, participants made a similarity rating for the lineup member on a Likert scale ranging from 1 (not similar at all) to 7 (completely similar). These similarity ratings were reverse-scored to obtain appearance change ratings. The video loops and photos of the four culprits used to obtain these similarity ratings can be viewed at http://www.psychology.iastate.edu/faculty/gwells/appearance_change_clips.html. In addition, each pre-test participant rated the similarity of each lineup filler to their respective targets in order to test for any differences between lineups in their fairness.

Participants and Design

Two hundred and eighty-nine undergraduate students participated in exchange for extra credit in their psychology courses. Participants were randomly assigned to either receive the appearance-change instruction or to not receive the instruction. Within each condition, each participant was shown four lineups, one for each of the four culprits (two males and two females) in the mock crime that they had previously witnessed. For each participant, two of these lineups were randomly assigned as culprit-present and the remaining two were culprit-absent, with the caveat that each participant saw one culprit-present male lineup and one culprit-present female lineup (and therefore each participant also saw one culprit-absent male lineup and one culprit-absent female lineup). The design of the experiment was a 2 (appearance-change instruction versus control) X 2 (present lineup versus absent lineup) factorial design, with appearance-change instruction as a between-subjects factor and lineup type as a within-subjects factor. Because each participant responded to four lineups, it is possible to think of this experiment as four replications of this 2 X 2 design. Sessions involved either one or two participants per session, each being placed in a separate room for the duration of the experiment.

Materials

Mock crime. Participants were shown on a computer a video of a mock crime that lasted for approximately five minutes. The video showed four individuals drive up to a locked garage, break the lock with bolt-cutters, enter the garage, root through the various items contained within the garage, take some of those items to their car, get scared off by the sound of someone coming, and drive off. During the video, each individual is in view for approximately 120 seconds. Each individual's face is clearly seen in close-up shots for approximately 15 to 20 seconds.

Lineups. Lineups were constructed from digital student identification card photographs, which are typically taken of students at the beginning of their freshman year. Because the four culprits in our mock crime were university students at different stages in their undergraduate education, they naturally varied in how much they had changed appearance since they had had their student identification card picture taken. Fillers (which were selected from a bank of digital student identification card photographs) were chosen based on their match to the culprit on the characteristics of gender, skin tone, approximate age, hair color, hair length, and facial hair. Six fillers were chosen per culprit. Culprit-absent lineups were composed of these six fillers whereas culprit-present lineups were composed of five of the fillers plus the culprit. Six different culprit-present lineups were created such that the culprit replaced a different one of the six fillers in each version. Participants viewing a culprit-present lineup were randomly shown one of these six versions. Thus, seven lineups (one culprit-absent plus six culprit-present) were created for each culprit. All photographs were high quality color photos and had similar backgrounds. Individuals in the photographs were shown straight-on from the shoulders up. All six lineup members were shown simultaneously on a computer screen in a 3(across) X 2(down) array. Each photograph in the lineup measured approximately 3 inches by 3.5 inches.

Procedure

Participants were initially told that they would be watching a video of people, and that we were interested in their impressions of some of the people in the video. Participants were informed that they would be seeing four individuals, and were given a brief description of each individual. Specifically, participants were told that they would see a dark-haired male, a light-haired male with glasses, a dark-haired female, and a light-haired female. Each participant was led into an individual cubicle and seated in front of a computer, where they were shown the mock crime. Following the mock crime, another video played of an experimenter informing the participants that they had just witnessed a crime and that they were now eyewitnesses. They were told that they would be shown four lineups, one for each of the culprits in the mock crime. Participants were told that each of the four lineups they were about to see would be referred to by a short description of the culprit on whom the lineup was based (i.e., dark-haired male, light-haired male with glasses, dark-haired female, light-haired female). Participants were given instructions on how to respond to the lineups, which involved pressing a key on the keyboard that was labeled with a number that corresponded to a lineup member's position in the lineup, or by pressing a key labeled "not there." Participants were told that for each of the four lineups they were about to see, the actual criminal may or may not be in the lineup. Finally, participants in the appearance-change instruction condition were further told that they should "keep in mind that the culprit may not look exactly the same in the lineup as he or she did during the crime. For example, the culprit may have changed hairstyle, or may be wearing different make-up, or may look either younger or older during the crime, or may look different in some other respect." Participants in the control condition were not given this last instruction.

Participants then viewed the four lineups in randomized order. Before every individual lineup, participants were informed as to which lineup they were about to see (dark-haired male, etc.), were reminded via written instructions that the culprit may or may not be present in the lineup and, for the appearance change participants, that the culprit "may not look exactly like he [she] did during the crime." Participants were shown each lineup as long as required to make a response (either an identification of a lineup member or a "not there" response). Once a response was made, participants were asked to state their confidence in their response to the previous lineup on a scale from 0%

to 100%, in 10% increments. After giving a confidence response, participants were automatically shown the next lineup. After participants had responded to all lineups, they were instructed to return to the main room where they were debriefed. All participants were asked if they recognized any of the criminals from the mock crime. Data from participants who recognized any of the criminals were excluded from analysis.

Results

Overview

We first analyzed the similarity ratings to determine whether appearance change differed across the four culprits, and whether lineup fairness differed across the four culprits. Then, we analyzed overall (collapsed across targets) choosing rates to find out if the appearance-change instruction affected choosing rates and, if so, to see if this effect interacted with the presence versus absence of the culprit. Next, we looked within the culprit-present conditions to see if the appearance-change instruction increased rates of identification of the culprits overall (collapsed over culprits) and for each individual culprit. Then, we analyzed the *proportion* of all identifications in culprit-present lineups that were of the target, both overall and for each lineup. [This proportion analysis is different from the previous analysis because it controls for any differences in choosing rates in the culprit-present conditions.] Then, we analyzed filler identification rates for culprit-present lineups, both overall and for each lineup, to see if they were affected by the appearance-change instruction. Next, we analyzed the choosing rates in culprit-absent conditions, both overall and for each lineup, as a function of the appearance-change instruction. Then, we calculated the diagnosticity of the lineups as a function of the appearance-change instruction. Finally, we analyzed the effects of the appearance-change instruction on confidence and on response latency.

Because each participant saw two culprit-present lineups and two culprit-absent lineups, each participant initially obtained a “culprit identification score” and a “filler identification score” of 0, 1, or 2 for culprit-present lineups, and a “false alarm score” of 0, 1, or 2 for culprit-absent lineups, which corresponded to the total number of culprit identifications, filler identifications, and false alarms, respectively, made by the participant. Each participant also obtained a “culprit-present choosing score” and a “culprit-absent choosing score” of 0, 1, or 2 that corresponded to the number of identifications made by that participant from culprit-present and culprit-absent lineups, respectively[2]. All scores were then divided by 2 (thus giving each participant a final score of 0, .5, or 1 for all overall measures) that corresponded to the number of actual identification responses divided by the number of possible identification responses. For example, imagine that Participant A made one culprit identification and one filler identification from the two culprit-present lineups seen, and did not make any identifications from the two culprit-absent lineups seen. Participant A would have obtained a culprit identification score of .5 (1/2), a filler identification score of .5 (1/2), a false alarm score of 0 (0/2), a culprit-present choosing score of 1 (2/2), and a culprit-absent choosing score of 0 (0/2). T-tests were performed on these scores for all overall analyses. Mean scores for culprit-present identification responses can be seen in the “Overall” column of Table 1. Mean scores for culprit-absent identification responses can be seen in the “Overall” column of Table 2. Tests of differences between proportions were used for all individual-culprit analyses.

A proper test of the appearance-change instruction requires that control participants be able to respond to the lineups at above-chance levels. As a manipulation check, culprit identification rates from target-present lineups were compared to the respective average false identification rates from target-absent lineups in the no instruction condition. If participants were responding at chance levels, then culprit identification rates should be equal to average false identification rates. Because there was no a priori innocent suspect, the average false identification rate is the total false identification rate divided by six (the number of lineup members). Collapsed across criminals, the culprit identification rate was significantly higher than the average false identification rate, $z = 8.45$, $p < .001$. At the level of the individual criminals, directional one-tailed tests indicated that culprit identification rates were higher than the respective average false identification rates for each of the four culprits (Culprit 1, $z = 8.06$, $p < .001$, Culprit 2, $z = 4.20$, $p < .001$, Culprit 3, $z = 1.29$, $p = .05$, and Culprit 4, $z = 2.48$, $p = .004$).

Appearance Change Analyses

The four culprits' similarity ratings, which were obtained during pre-testing, were reverse-coded to attain an appearance-change score, and this score was compared across culprits. A Bonferroni correction was made for the number of pairwise comparisons made. The four culprits will be referred to throughout this section as Culprit 1, Culprit 2, Culprit 3, and Culprit 4, in order of increasing appearance change. Culprit 1 ($M = 1.64$) was rated as having undergone significantly less appearance change than Culprit 2 ($M = 2.62$), $t(60) = 4.94$, $p < .001$, $d = .64$. Culprit 2 was rated as having undergone significantly less appearance change than Culprit 3 ($M = 3.50$), $t(61) = 3.44$, $p = .001$, $d =$

.44, and Culprit 4 (*M* = 3.71), *t*(62) = 4.14, *p* < .001, *d* = .53. Culprit 3 and Culprit 4 did not differ significantly from each other in terms of the degree of appearance change, *t*(61) = .80, *p* = .37, *d* = .10. The significant differences between the various culprits indicate that our appearance change manipulation was successful. Readers can view the video loops and still photos of the four culprits at http://www.psychology.iastate.edu/faculty/gwells/appearance_change_clips.html.

Relative Fairness of Fillers Across Lineups

In order to ascertain the relative fairness of the various lineups, the pretest participants' ratings of similarity of the fillers to targets were used. Because the current study did not have an a priori innocent suspect, an average filler similarity score was calculated for each of the four culprits. If lineup fairness differed across target-absent lineups, then the average filler similarity scores should differ as well. A one-way ANOVA failed to show any significant differences in average filler similarity scores between the target-absent lineups for Culprit 1 (*M* = 2.97), Culprit 2 (*M* = 2.82), Culprit 3 (*M* = 2.83), and Culprit 4 (*M* = 2.92), *F*(3, 180) = .86, *p* = .46. Thus, lineup fairness did not differ across the four criminals' target-absent lineups.

Choosing Rates

Using choosing scores as the units of analyses, a 2 (appearance-change instruction) X 2 (present versus absent lineup) mixed ANOVA found that the appearance-change instruction significantly increased choosing rates, *F*(1, 287) = 34.5, *p* < .001, Cohen's *f* = .35. This relationship was not moderated by a significant interaction, *F*(1, 287) = .08, *p* = .78, Cohen's *f* = .00, indicating that the appearance-change instruction increased choosing rates equally from culprit-present and culprit-absent lineups (*M* = .60

for instructions; *M* = .45 for no instructions). Upon first glance, this increase in choosing rates from both culprit-present and culprit-absent lineups, coupled with a lack of interaction, does not bode well for the usefulness of the appearance-change instruction. However, it is possible that the increase in choosing rates from culprit-present lineups reflects primarily an increase in culprit, as opposed to filler, identifications, whereas the increase in choosing rates from culprit-absent lineups is distributed across all lineup members. If this is true, then under the assumption of a single-suspect model (see footnote 1), most of the increase in culprit-absent identifications would result in the identification of fillers (who are known to be innocent of the crime), as opposed to the identification of the innocent suspect. Therefore, the appearance-change instruction might still be useful if its net effect is to increase culprit identifications at a greater rate than filler identifications and innocent suspect identifications.

Culprit Identification Rates From Culprit-Present Lineups

Collapsed across culprits, the appearance-change instruction did not significantly affect the mean culprit identification score ($M = .30$ for instructions; $M = .30$ for no instructions), $t(287) = .03$, $p = .98$, $d = .00$. At the level of each lineup, the appearance-change instruction did not significantly affect the culprit identification rates for Culprit 1, Culprit 2, or Culprit 3, $z = -.26$, $p = .40$ (Culprit 1), $z = -.97$, $p = .17$ (Culprit 2), $z = -.24$, $p = .41$ (Culprit 3). The appearance-change instruction increased the culprit identification rate for Culprit 4 to a marginally-significant extent, $z = 1.59$, $p = .06$.

Our next analysis examined culprit identification rates as a *proportion* of all identification attempts in the culprit-present lineups. Because each participant saw two culprit-present lineups, each participant could have attempted up to two identifications and could have identified the actual culprit up to two times (with the obvious constraint that a witness's actual culprit identifications could not exceed his or her attempted identifications). Thus, for the overall analysis, each participant obtained a score of 0, .5, or 1 that corresponded to the proportion of culprit-present identification attempts that were of the culprit. Collapsed across culprits, the appearance-change instruction significantly diminished the proportion of all identifications that were of the culprit (.51 for instructions; .66 for no instructions), $t(218) = 2.46$, $p = .01$, $d = .34$. Examining culprits individually, the appearance-change instruction had no significant effect on the proportion of culprit-present identifications that were of Culprit 1 (.80 for instructions; .80 for no instructions), $z = -.06$, $p = .48$, Culprit 3 (.20 for instructions; .35 for no instructions), $z = -.71$, $p = .24$, or Culprit 4 (.53 for instructions; .58 for no instructions), $z = -.09$, $p = .46$. The appearance-change instruction did, however, significantly reduce the proportion of culprit-present identifications that were of Culprit 2 (.32 for instructions, .69 for no instructions), $z = -1.65$, $p = .050$. The appearance-change instruction thus failed to increase the proportion of all identification attempts that were of the culprit.

Filler Identification Rates From Culprit-Present Lineups

Collapsed across culprits, the appearance-change instruction significantly increased the mean filler identification score ($M = .30$ for instructions; $M = .15$ for no instructions), $t(287) = 3.94$, $p < .001$, $d = .46$. At the level of the individual lineups, the appearance-change instruction did not significantly increase the filler identification rate for Culprit 1, $z = -.13$, $p = .45$. However, the instructions did significantly increase filler identification rates for the remaining culprits, $z = 3.04$, $p = .001$ (Culprit 2), $z = 1.81$, $p = .03$ (Culprit 3), $z = 1.84$, $p = .03$ (Culprit 4).

False Alarm Rates from Culprit-Absent Lineups

Collapsed across culprits, the appearance-change instruction significantly increased the mean false alarm score ($M = .51$ for instructions; $M = .31$ for no instructions), $t(287) = 5.02$, $p < .001$, $d = .59$. At the level of the individual lineups, the appearance-change instruction did not significantly increase false alarms for Culprit 1, $z = .55$, $p = .29$. However, the appearance-change instruction did significantly increase false alarms for the remaining culprits, $z = 2.83$, $p = .002$ (Culprit 2), $z = 2.92$, $p = .002$ (Culprit 3), $z = 1.71$, $p = .04$ (Culprit 4).

Diagnosticity

Diagnosticity is a statistic that can be used to estimate the probative value of evidence and is useful for comparing one lineup procedure to another to find out which procedure yields the greater probative value. First introduced by Wells and Lindsay (1980), the diagnosticity of the identification of a suspect is simply an expression of how much more likely it is that the suspect, if identified, is the culprit rather than an innocent person. Diagnosticity estimates require either the a-priori specification of an innocent suspect from the culprit-absent lineups or the use of an average identification rate per lineup member from the culprit-absent lineups. Because the current work rotated the suspect versus filler role among all culprit-absent members, we used the average rate per lineup member (percent of identifications divided by 6) for the culprit-absent lineups for all estimates of diagnosticity. Extensive treatments of the diagnosticity index can be found elsewhere (Wells & Turtle, 1986; Wells & Olson, 2002).

Collapsed over culprits, diagnosticity of suspect identifications was 5.78 without the appearance-change instruction and 3.53 with the appearance-change instruction. This means that an identification of a suspect was 5.78 times as likely to be the culprit than to be an innocent suspect when the appearance-change instruction was not given and only 3.53 times as likely to be the culprit than to be an innocent suspect when the appearance-change instruction was given. Based on an interaction test among proportions (see Wells & Olson, 2002), the lower diagnosticity of suspect identifications associated with the appearance-change instruction was not statistically significant, $z = 1.11$, $p = .15$.

At the level of each culprit, the diagnosticity of suspect identifications for those in the no-appearance-change instruction and appearance-change instruction conditions, respectively, were 7.51 and 6.55 for Culprit 1, 7.58 and 2.17 for Culprit 2, 2.08 and .93 for Culprit 3, and 4.93 and 4.96 for Culprit 4. The only statistically significant difference was for Culprit 2 for which diagnosticity was greater without rather than with the appearance-change instruction, $z = 1.74$, $p = .04$.

Confidence and response latency

Participants who were given the appearance-change instruction reported significantly lower average confidence in their identification decisions than participants who were not given the appearance-change instruction ($M = 63.1$ for instructions; $M = 66.6$ for no instructions), $t(287) = 2.01$, $p = .045$, $d = .12$. The appearance-change instruction significantly increased participants' average response latency, from presentation of lineup to identification response ($M = 19.3$ s for instructions; $M = 16.0$ s for no instructions), $t(287) = 3.40$, $p = .001$, $d = .20$.

Discussion

Results of various analyses clearly show the lack of beneficial effects of the appearance-change instruction. In fact, the appearance-change instruction was actually detrimental to the identification process. The instruction significantly increased identifications from target-absent lineups from 31% without instructions to 51% with instructions. It should be noted that this is not quite as alarming as it first appears because, under the assumptions of a fair lineup and the single-suspect model, we would expect only 1/6th of these identifications to be of an innocent suspect, the remainder being known-innocent fillers. Nevertheless, the rate of mistaken identifications of an innocent suspect was inflated by over 50% for participant witnesses who received the appearance-change instruction.

Interestingly, this increase in false identifications was not coupled with an overall increase in culprit identifications. In fact, of the four culprits, the appearance-change instruction only significantly increased identifications of Culprit 4 (albeit marginally so), an increase that was not significantly different from the increase in false identifications when culprit 4 was not present in the lineup. Additionally, the appearance-change instruction significantly reduced the overall proportion of identifications from target-present lineups that were of the criminal. In fact, although the appearance-change instruction increased filler identifications overall, it did not increase culprit identifications overall, a result that was unexpected and surprising.

Why did the instruction not increase hits, as we had hypothesized? Clearly, witnesses vary from one to another with respect to how good a memory they had of the actual criminal for a variety of reasons (individual differences in encoding, how much attention was paid during the crime, etc.). Assume now that when the appearance-change instruction was not given prior to being shown a culprit-present lineup, those witnesses who had a better memory of the criminal tended to identify a lineup member whereas those witnesses who had a poorer memory of the criminal tended (appropriately) to reject the lineup. The increase in choosing resulting from the appearance-change instruction would thus reflect primarily an increase in choosing among those witnesses who had a poor memory of the criminal. Given that their memory for the culprit was poor, their identifications should be more or less randomly distributed across lineup members. Because there were five fillers per lineup, the increased identification rate of any given filler is the total increase in filler identifications (14.4%) divided by 5, or 2.8%. Assuming a roughly random distribution of identifications, culprit identifications should therefore have increased about 2.8% as a result of the appearance-change instruction. The observed increase in culprit identifications (.1%) is within the margin of error of this hypothesized increase. Thus, although surprising at first glance, the observed results are consistent with the assumption that the appearance-change instruction specifically leads those witnesses who have poor memories to make identifications.

This study provided four tests of the appearance-change instruction, one for each of the four culprits. Importantly, the appearance-change instruction did not increase identification accuracy for any of the culprits, despite the range of appearance change undergone by the four culprits. In fact, the difference in appearance-change scores between the least changed and the second-least changed was .64 standard deviations and the difference between the second-least changed and the most changed was an additional .53 standard deviations. It is thus unlikely that the lack of beneficial effects of the appearance-change instruction is a result of using a restricted range of appearance change of the culprits.

Recall that three processes (a selective increase in ecphoric similarity, a general increase in ecphoric similarity, and a lowering of decision criterion) were proposed that could explain potential appearance-change instruction effects. What light can the observed results now shed on these processes? The appearance-change instruction increased choosing rates for both target-present and target-absent lineups with no increase in accuracy. Thus, the current study's findings clearly are not consistent with a selective increase in ecphoric similarity, which should have resulted in the appearance-change instruction producing a greater increase in culprit identifications than in filler identifications. These results are consistent, however, with both a general increase in ecphoric similarity and a lowered decision criterion. Are there any data that can somehow distinguish between these two theoretical explanations?

Apart from actual identifications, two other measures were also collected, namely confidence and response latency. Unfortunately, neither of these measures allow us to differentiate between a general increase in ecphoric similarity and a lowered decision criterion, as the observed results can be explained by both hypotheses. For instance, participants who received the appearance-change instruction reported having significantly lower confidence in their identifications than participants who did not receive the appearance-change instruction. This finding could be the result of a general increase in ecphoric similarity: Mental mutations may lead multiple lineup members to be able to fit one's memory of the culprit, thus creating more competition for an identification, resulting in lowered confidence. Alternatively, the finding could be the result of a lowered decision criterion: More lineup members would surpass the newly-lowered criterion, thus creating more competition for an identification, resulting in lowered confidence. Note, however, that this lower confidence associated with the appearance-change instruction is inconsistent with a selective increase in ecphoric similarity, for reasons described earlier. Participants who received the appearance-change instruction also showed a significantly longer response latency than participants who did not receive the appearance-change instruction, as we had predicted. This finding could be the result of a general increase in ecphoric similarity: Mental mutations are cognitive processes that take time, and thus encouraging these mutations should increase response latency. Or perhaps it is the result of a lowered decision criterion: If more lineup members surpass one's newly-lowered criterion, it may take more time for a witness to determine which of those lineup members should be identified.

Although we are unable to distinguish between a general increase in ecphoric similarity and a lowered decision criterion, both the identification data and the confidence data are inconsistent with the existence of a selective increase in ecphoric similarity. Thus, regardless of the exact theoretical underpinning of the appearance-change instruction effects, the data are not favorable towards the use of the appearance-change instruction.

These findings raise interesting questions about how the appearance-change instruction recommendation came to be a part of the Department of Justice Guide (Technical Working Group, 1999) and also more general concerns about the role of psychological science in legal reform. A general account of the processes involved in developing the Guide (but no discussion of the appearance-change instruction) is described elsewhere by the six researchers who were part of the Technical Working Group that authored the Guide (Wells, Malpass, Lindsay, Fisher, Turtle, & Fulero, 2000).

As for the appearance-change instruction, this was a recommendation that was made by the law enforcement members (n = 17) of the Technical Working Group who argued that this type of instruction was already used in many of their jurisdictions. The general view of the researchers was that their role was to inform the Technical Working Group of all relevant data and to shape the recommendations toward procedures that could be defended based on those data. In the case of the appearance-change instruction, there were no relevant data. Accordingly, the researchers remained largely silent on the appearance-change instruction recommendation that came from law enforcement. In hindsight, it might have been wise for the researchers to question the recommendation by discussing the possibility that the appearance-change instruction could have negative effects (e.g., by lowering the decision criterion of eyewitnesses). But the general rule, which the researchers imposed upon themselves, was to argue primarily only those matters for which there were relevant data or a general principle that was clearly applicable to the question.

This raises a more general and recurring question about the role of psychological scientists in shaping legal opinions and policies. Should psychological scientists restrict themselves to offering opinions to policy makers only when there are directly relevant data? At what point might the psychological scientists have a good enough conceptual understanding of the relevant processes that, even in the absence of directly relevant data, the psychological scientist can and should make extrapolations? In some cases, the answer seems obvious. For example, there are no data on what happens when the photo of a suspect in a photo-lineup is in color and the filler photos are black and white. Nevertheless, few eyewitness researchers would hesitate to offer the opinion that the lineup was biased against the suspect. In this case, however, eyewitness researchers have conducted enough research to extract a more general principle that the suspect's photo should not stand out in a photo-lineup. But what if there were no direct data and also the situation did not clearly fall under a general principle? The issue is made complex by the fact that policy-makers will nonetheless make decisions regarding procedures even if the psychological scientist remains silent, as happened in the case of the appearance-change instruction. Should the psychological scientist remain silent on the matter? We do not know the answer to that question, but we are sure that there is a general consensus among researchers that, given a lack of scientific knowledge about an issue, it is desirable to return to the lab to collect the data relevant to addressing the issue.

We are in no way suggesting that the current research has definitively resolved the question of whether the appearance-change instruction is detrimental. No single study, regardless of how soundly it was designed and conducted, should be considered conclusive. In the case of the appearance-change instruction, for example, there could be alternative wordings of the instruction that might prove beneficial. Also, perhaps there is some way to determine a-priori whether an appearance-change instruction should be given or not. We note, for example, that the appearance-change instruction increased hit rates for the culprit that had the most appearance change (Culprit 4) in the current research. Perhaps the appearance-change instruction should only be given when the criminal has undergone significant appearance change. But how would investigators know a-priori whether the culprit had undergone significant appearance change in an actual case? One possibility is to examine discrepancies between the verbal description that the eyewitness gave of the culprit and the physical appearance of the suspect.

However, we know that verbal descriptions of culprits commonly vary from the actual physical characteristics of the culprit, so the discrepancy could be attributable to a bad description rather than appearance change. Furthermore, it is always possible that the discrepancy between the verbal description of the culprit and the physical appearance of the suspect is due to the fact that the suspect is not the culprit. Hence, a discrepancy between the description of the culprit and the physical appearance of the suspect is a questionable basis for assuming appearance change.

Another situation in which investigators might assume that there has been appearance change is when it is known that the suspect has in fact changed appearance since the time of the crime. For example, investigators might question friends or relatives of the accused and learn that there has been appearance change (e.g., suspect added or removed facial hair, gained or lost weight). Again, however, this presumes that the suspect is in fact the culprit. Recall that when the suspect was not the culprit (the culprit-absent lineups), the appearance-change instruction increased mistaken identifications. Hence, we do not see a practical way to tailor the appearance-change instruction to specific situations at this time.

Whether the appearance-change instruction is ultimately considered a bad idea or not awaits more research, but these data raise serious doubts about the wisdom of the appearance-change instruction. On the more positive side, this research highlights the importance of one of the other recommendations made in the Guide. Specifically, the observed differences in hit rates across criminals who underwent varying degrees of appearance change underscores the importance of trying to make sure that the physical appearance of a suspect in a photo lineup is a reasonable representation of how that suspect appeared at the approximate time of the crime. The Guide makes precisely that recommendation when stating that “If multiple photos of the suspect are reasonably available to the investigator, select a photo that resembles the suspect[‘s] … appearance at the time of the crime” (p. 29). This recommendation is especially important to the extent that the appearance-change instruction ultimately proves to have detrimental effects.

References

Behrman, B.W., & Davey, S.L. (2001). Eyewitness identification in actual criminal cases: An archival analysis. *Law and Human Behavior, 25*, 475-491.

Charman, S. D. and Wells, G. L. (in press). Applied lineup theory. In D. Read et al. (Eds.) *Handbook of Eyewitness Psychology*. Mahwah, NJ: Lawrence Erlbaum Associates.

Clark, S.E. (2003). A memory and decision model for eyewitness identification. *Applied Cognitive Psychology, 17*, 629-654.

Clark, S. E. (2005). A Re-examination of the effects of biased lineup instructions in eyewitness identification. *Law and Human Behavior, 29,* 395-424.

Cutler, B.L., & Penrod, S.D. (1995). *Mistaken identification: The eyewitness, psychology, and the law*. New York: Cambridge University Press.

Ebbesen, E., & Flowe, S. (2002). *Simultaneous versus sequential lineups: What do we really know?* Retrieved July 29, 2002 from http://www.psy.ucsd.edu/%7eeebbesen/SimSeq.htm

Kahneman, D. & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.) *Judgment under uncertainty: Heuristics and biases* (p. 201-210). New York: Cambridge University Press.

Kassin, S.M., Tubb, V.A., Hosch, H.M., & Memon, A. (2001). On the "general acceptance" of eyewitness testimony research. *American Psychologist, 56*, 405-416

Lindsay, R.C.L., & Wells, G.L. (1980). What price justice? Exploring the relationship between lineup fairness and identification accuracy. *Law and Human Behavior, 4*, 303-314.

Malpass, R.S., & Devine, P.G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology, 66*, 482-489.

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. Memory & Cognition, 33, 783-792.Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many disguises. *Review of General Psychology, 2*, 175-220.

Shapiro, P. N. & Penrod, S. D. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100,* 139-156.

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

Tulving, E. (1981). Similarity relations in recognition. *Journal of Verbal Learning and Verbal Behavior, 20*, 479-496.

Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology, 17*, 969-993.

Wells, G.L. (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology, 14*, 89-103.

Wells, G.L., Malpass, R.S., Lindsay, R.C.L., Fisher, R.P., Turtle, J.W., & Fulero, S.M. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist, 55,* 581-598.

Wells, G. L. & Olson, E. (2002). Eyewitness identification: Information gain from incriminating and exonerating behaviors. *Journal of Experimental Psychology: Applied, 8*, 155-167.

Wells, G.L., Small, M., Penrod, S., Malpass, R.S., Fulero, S.M., & Brimacombe, C.A.E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads, *Law and Human Behavior, 22*, 603-647.

Wells, G.L., & Turtle, J.W. (1986). Eyewitness identification: The importance of lineup models. *Psychological Bulletin, 99*, 320-329.]

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

Wright, D.B., & McDaid, A.T. (1996). Comparing system and estimator variables using data from real lineups. *Applied Cognitive Psychology, 10*, 75-84.

Footnotes

1. A single-suspect lineup means that one member of the lineup is a suspect (who might or might not be the culprit) and the remaining members are known-innocent fillers. Some actual lineups have included all suspects and no fillers (the all-suspect lineup model) and others have included multiple suspects and multiple fillers (mixed lineup models). These different models for the lineup have been discussed and analyzed extensively by Wells and Turtle (1986) who concluded that the mixed and all-suspect models fail miserably at controlling mistaken identification rates. Both the Department of Justice Guide (Technical Working Group, 1999) and the American Psychology-Law Society lineups review paper (Wells et al., 1998) strongly endorse only the single-suspect model for lineups. Accordingly, the current research assumes the single-suspect model. Among the implications of the single-suspect model relevant to the current research are: (a) identifications of innocent suspects cannot occur with culprit-present lineups because everyone other than the culprit is a known-innocent filler, not a suspect, (b) only one member of a culprit-absent lineup is an innocent suspect, the remainder being known-innocent fillers, and (c) under the assumption of a fair lineup, 1/N (where N is the number of lineup members) of the choices from a culprit-absent lineup are considered "forensically-relevant" false alarms.

2. Note that because every identification from a culprit-absent lineup was necessarily a false alarm in this study, participants' culprit-absent choosing scores are identical to their false alarm scores.

Table 1

Proportion of culprit identifications, filler identifications, and "not there" responses from culprit-present lineups under conditions of appearance-change instructions and no appearance-change instructions

	Culprit				
Lineup response	1	2	3	4	Overall
Instructions					
Culprit identification	.667	.181	.097	.264	.302
Filler identification	.166	.388	.389	.236	.295
"Not there"	.167	.431	.514	.500	.403
No instructions					
Culprit identification	.694	.265	.113	.133	.301
Filler identification	.177	.121	.210	.096	.151
"Not there"	.129	.614	.677	.771	.548

Table 2

Proportion of false identifications and correct rejections from culprit-absent lineups under conditions of appearance-change instructions and no appearance-change instructions

	Culprit				
Lineup response	1	2	3	4	Overall
Instructions					
False identification	.611	.500	.625	.319	.514
Correct rejection	.389	.500	.375	.681	.486
No instructions					
False identification	.554	.210	.325	.161	.313
Correct rejection	.446	.790	.675	.839	.688