Policy and Procedure Recommendations for the Collection and Preservation of Eyewitness Identification Evidence

Gary L. Wells
Iowa State University

Margaret Bull Kovera
John Jay College and the Graduate Center, City University of New York

Amy Bradfield Douglass
Bates College

Neil Brewer
Flinders University

Chris Meissner
Iowa State University

John T. Wixted
University of California – San Diego
Abstract

The American Psychology/Law Society and Division 41 of the American Psychological Association appointed a subcommittee to update the influential 1998 scientific review paper on guidelines for eyewitness identification procedures. This new review resulted in a reinstatement of the four original recommendations, which concerned the selection of lineup fillers, pre-lineup instructions to witnesses, the use of double-blind procedures, and collection of a confidence statement. An additional five recommendations were made in this new review, which concern (1) the need for reasonable-suspicion before conducting an identification procedure, (2) the conductance of a pre-lineup interview of the witness, (3) video-recording of the entire procedure, (4) avoiding repeated identification attempts with the same witness and same suspect, and (5) avoiding the use of showups when possible and improving how showups are conducted when they are necessary.
Policy and Procedure Recommendations for the Collection and Preservation of Eyewitness Identification Evidence

The identification of criminal suspects by eyewitnesses continues to be a staple form of evidence used by the justice system to help establish the identity of crime perpetrators. Despite the increased use of surveillance recording devices as well as the advent of forensic DNA testing and other forensic evidence techniques, eyewitness identification evidence from victims and bystanders remains central to solving crimes.

Over the last few decades, however, serious concerns have been raised about the potential unreliability of eyewitness identification evidence. There have been two primary forces that have helped shape this concern about the reliability of eyewitness identifications. First, psychological scientists began developing programmatic lab-based experimental research starting in the mid to late 1970s that focused on eyewitness identification. Early in this research it became readily apparent that certain variables that are under control of the criminal justice system can dramatically inflate rates of mistaken identification; these variables were called *system variables* (Wells, 1978). An instruction given to eyewitnesses prior to viewing a lineup, for example, is a system variable because it influences the reliability of the eyewitness’s identification, and is under the control of the legal system’s policies and procedures for administering lineups (Malpass & Devine, 1981). By the mid 1990s, psychological scientists had published hundreds of lab-based experiments showing that mistaken identification rates can be very high under certain conditions and had identified some of the more problematic sets of conditions that can lead to such errors in real-world circumstances (e.g., see Cutler & Penrod, 1995).
The second major force propelling a strong awareness of the potential fallibilities of eyewitness identification was the advent of forensic DNA testing in the mid and late 1990s. Although forensic DNA testing was conceived primarily as a tool to incriminate the guilty, the exonerating powers of forensic DNA testing quickly became evident. Starting in the 1990s, DNA began to be used to test claims of innocence in selected post-conviction cases and a cascade of exonerations of innocent people began to unfold. In a report commissioned by the U.S. Department of Justice examining the first 30 exonerations, it was quite clear that the vast majority of these cases involved mistaken eyewitness identifications (Conners et al., 1996). Although legal scholars had described a number of cases of innocent people being convicted in the United States based on mistaken identification before the development of DNA testing (e.g., see Borchard, 1932; Brandon & Davies, 1973; Frank & Frank, 1957; Huff, Ratner & Sagarin, 1986), the DNA exoneration cases quickly outnumbered previously-known cases of convictions of the innocent. Moreover, the pre-DNA cases tended to be less definitive about whether the person was guilty, often just indicating “probable innocence” or “legal innocence” rather than the more definitive “actual innocence” characterization that accompanied DNA-based exonerations (Scheck, Neufeld, & Dwyer, 2000).

The development of a scientific literature on eyewitness identification and the use of forensic DNA testing to uncover mistaken eyewitness identifications were a powerful combination of forces. In 1996, the Executive Committee of the American Psychology-Law Society (APLS), the primary scholarly organization for eyewitness identification researchers, appointed a committee to draft a scientific review paper on recommendations for how to collect eyewitness identification evidence. The result of that review, vetted by the membership of APLS and subjected to peer review, was published in Law and Human Behavior (Wells, Small, Penrod,
Malpass, Fulero, & Brimacombe, 1998). That article, commonly referred to as the “white paper on lineups,” is the forerunner to the current article. In effect, the current article is an update on the original scientific review paper on lineups and, like the original scientific review paper, also represents the official position of the American Psychology-Law Society on these issues. The 1998 scientific review article was the first set of science-based recommendations regarding how to conduct lineups that was endorsed by a scientific society. In addition, the 1998 scientific review paper played an important role in subsequent developments. For example, it was the model on which the U.S. Department of Justice made its recommendations in its 1999 guide for law enforcement on collecting and preserving eyewitness evidence that was mailed to all of the more that 17,000 law enforcement agencies in the U. S. (Technical Working Group, 1999).

Moreover, the 1998 article was a model on which the Department of Justice in New Jersey created the first statewide guidelines on eyewitness identification that law enforcement must follow. It was also the first peer-reviewed publication to include a description of the unfolding DNA exoneration cases (40 at that time) and what these DNA cases might tell us about eyewitness identification evidence. Finally, the 1998 scientific review paper has been used as supporting written material in countless seminars and workshops of lawyers, police, prosecutors, and judges as well as police training academies across the U.S.

**The Need for an Updated Scientific Review Paper on Lineups**

Today, our understanding of eyewitness identification has matured well beyond where it was when the previous scientific review paper was published. Clearly, experimental lab studies have grown immensely in number, quality, and breadth over the last 20 years. However, the general methodology of eyewitness identification studies has largely remained the same. People are exposed to a simulated crime, usually via video, and the researchers know the true identity of
the culprit. These participant-witnesses are then shown a lineup, typically a photo-lineup, in which the culprit’s photo is embedded among filler photos or the culprit’s photo is absent from the array and replaced with the photo of an innocent person. Using this basic paradigm, researchers then systematically manipulate variables, such as the view that the witness had of the culprit, the similarity of the fillers to the culprit, the instructions given to the eyewitness prior to viewing, suggestive behaviors of the lineup administrator, and so on, to see how those variables affect the responses of the eyewitnesses. This experimental lab methodology has a number of strengths that arise from the fact that “ground truth” is known (i.e., the researchers know which person is the actual “perpetrator”) and from the use of random assignment to conditions that permit inferences about the cause of any effects observed in responses.

In addition to the growth of experimental lab studies since 1998, a second type of research methodology has emerged that examines eyewitness identifications from lineups by actual witnesses (victims and bystanders) to crimes. These “field” studies have clear limits, in large part because of the difficulty of establishing ground truth for identifications of suspects, but also because of multi-collinearity issues (e.g., see Horry, Halford, Brewer, Milne, & Bull, 2014). However, there is an important aspect of these field study data that helps counter some of the criticisms of lab-based experiments, in particular the criticism that participants in lab experiments make a lot of mistaken identifications only because the witnessed events are not real and the consequences of mistaken identification are trivial. The argument is that real witnesses to serious crimes would not be so careless. Hence, the argument is that lab studies might overestimate the eyewitness identification problem. The results of peer-reviewed field studies, however, show otherwise. In these field studies of lineup outcomes, the lineup contains “fillers” who are known to be innocent of the crime (such as people who were in prison at the time of the
EYEWITNESS IDENTIFICATION RECOMMENDATIONS

If the eyewitness identifies the suspect, it might or might not be an accurate identification because ground truth is not known with certainty in actual cases. But when an eyewitness identifies a filler, it is clearly a mistaken identification because these fillers are known to be innocent. A mistaken identification of a filler is low risk because it will not result in charges against that filler but a filler identification is a mistaken identification nevertheless.

How often do actual eyewitnesses in serious crime cases mistakenly identify a filler from a lineup? There are now 11 peer-reviewed published studies from which we can estimate filler identification rates based on 6,734 lineups. These field studies are from highly varied jurisdictions (e.g., California, Arizona, Texas, London, England) and a summary of these data is shown in Table 1. [Some studies reported data that included identifications by witnesses who knew the perpetrator (prior familiarity) and those data are excluded from Table 1.] For current purposes, two statistics of note from Table 1 speak to the question of whether actual witnesses to serious crimes are too cautious to make mistaken identifications at rates like those observed in lab experiments. First, nearly one of every four witnesses (23.7%) who was shown a lineup picked an innocent filler. Second, among those who made an identification (35.5% made no identification), over one third (36.8%) identified a known-innocent filler. These field study data, which were not available at the time of the 1998 scientific review paper, put to rest the argument that lab studies produce highly-inflated rates of mistaken identification compared to what would happen with actual eyewitnesses to serious crimes. Mistaken identifications to fillers occur frequently in both lab studies and field studies, though with much less confidence, on average, than suspect identifications.

In addition to a large growth in the size and quality of lab-based experiments as well as the addition of field studies of lineups in actual cases, DNA exoneration cases involving
mistaken eyewitness identification have continued to accumulate. As of this writing, there have been 248 people in the U.S. who were mistakenly identified (often by multiple witnesses), convicted, and serving prison sentences (some on death row) until forensic DNA testing proved their innocence and they were officially exonerated (Innocence Project, 2018). Approximately 70% of DNA exonerations are cases that involved mistaken eyewitness identification. Although the number of DNA exonerations represent only a small fraction of convictions, it is essential to note that these exonerees were the lucky few for whom DNA-rich trace evidence for the crime existed, was collected properly, and was preserved. Contrary to public perceptions based on television programs, perpetrators leave behind DNA-rich trace evidence in only a small fraction of cases. The largest category of convictions based on eyewitness identification evidence is robberies, and perpetrators of robbery almost never leave behind DNA-rich trace evidence that could exculpate a mistakenly identified person. For cases in which DNA did exist (primarily restricted to sexual assault cases), no one anticipated that forensic DNA testing would later develop and so for most cases the trace evidence was not collected, was not preserved properly (allowing it to deteriorate), was lost, or was destroyed by the time forensic DNA testing came along. Hence, DNA testing could only reverse a very small fraction of possible mistaken identifications. Indeed, eyewitness identification evidence is still heavily relied on today because DNA and other forms of definitive evidence remain extremely rare. Improving the reliability of eyewitness identification evidence therefore remains an important priority in preventing miscarriages of justice.

Focus of the Current Article

The current article, like the 1998 scientific review paper, is not a broad review of all issues in eyewitness identification. It is, instead, a focused examination of system variables in
eyewitness identification; factors that relate to the reliability of eyewitness identifications over which the justice system has (or can have) control. Hence, despite the fact that there are many powerful variables that affect the reliability of eyewitness identification evidence called estimator variables (e.g., same versus cross-race identifications, fear and stress during the witnessing of a crime, quality of view) and so on, such variables are not under control of the justice system and not the focus of this scientific review paper.

In some ways, the definition of system variables that is used today has broadened from its original definition. Originally, system variables in eyewitness identification referred to variables that influence the accuracy of eyewitness identifications over which the justice system has control (Wells, 1978). Over time, however, the definition of system variables has broadened somewhat to include factors under the control of the justice system that relate to (as opposed to influence) the accuracy of eyewitness identifications. Eyewitness confidence, for example, does not itself influence the accuracy of an eyewitness identification. However, eyewitness confidence is related to the accuracy of eyewitness identification and it is easily contaminated by events that are under the control of the criminal justice system (such as feedback from the lineup administrator). Accordingly, securing a confidence statement at the time of identification using a double-blind lineup administrator is a system variable (see Wilford and Wells, 2013 for a more extended discussion of this broader view of system variables).

In the current paper, other examples of this broader definition of system variables will become apparent. For example, one of the recommendations in this new scientific review paper is video-recording the entire identification procedure. Obviously, video-recording the identification procedure is not meant to increase the accuracy of eyewitness identifications. Instead, the purpose of the video-record recommendation is to secure a record that can help to
assess the quality of the identification and the procedure. That is, a video can shed light on the likely accuracy of eyewitness identifications, provide a record of the potential suggestiveness in procedures, and provide other potential benefits as discussed later in this article. Using the broader definition, video-recording qualifies as a system variable because it is under the control of the justice system to either video-record or not record, and the recording is relevant to the eyewitness accuracy problem.

A central issue in the development of a set of recommendations on policies and procedures in eyewitness identification is how to decide which recommendations are the most important ones to make and what criteria ought to be used to decide whether to include a recommendation. In an ideal world, we might say that we should recommend only those policies and procedures for which there is overwhelming scientific evidence. For most of our recommendations, there is a solid and specific body of scientific evidence to support the recommendation and we review that scientific evidence. In some cases, however, the recommendation is based primarily on reasonably well-established understandings of human memory and social influence, common sense, our general understanding of problems that we have observed in actual cases, or all three foundations. Consider again our recommendation that the entire identification procedure be video-recorded. We believe that this can have many benefits, including but not restricted to: prophylactic effects on potential suggestive behaviors by the lineup administrator, proof as to exactly what instructions were given to the eyewitness, information about how long it took the eyewitness to make an identification, and verbal and nonverbal records of the confidence of the eyewitness. We believe that the arguments favoring this recommendation are compelling. However, there has been little research on the issue of whether they serve a prophylactic function, how the videos should be used, on whom the camera
should be focused, and so on. Under a strict criterion of requiring “overwhelming scientific evidence,” the video-recording recommendation might not have made the cut. However, we believe that it would be irresponsible to not include a video-recording recommendation for reasons that will become clear in that section. For each of the recommendations that we make, we include justifications for the recommendation as well as discussion of any concerns or caveats.

**Live versus Photo-Lineups**

Throughout this article, we use the term lineup to refer to both live lineups and photo-lineups. We do not include separate sections on these two common procedures. This is because we know of no evidence to indicate that the principles governing photo-lineups and live lineups are different. In others words, every recommendation that we make about lineups applies equally to live and photographic displays.

Photo-lineups are far more common than live lineups in the U.S. (Police Executive Research Forum, 2013). Photo-lineups are sometimes called photo-spreads, photo-arrays, or photo-montages. Another common name for photo-lineups in many U.S. law enforcement circles is “six packs,” which refers to the most frequent size and arrangement of a photo-lineup in which the lineup contains six photos arranged in two rows of three. Live lineups, sometimes called corporeal lineups, are relatively rare compared to photo-lineups, likely owing to the greater difficulty, time, and expense of constructing and conducting live lineups relative to photographic lineups.

Even though the principles governing live versus photo-lineups are thought to be the same, questions have been raised as to whether performance overall might be better for live lineups than for photo-lineups. The *live-superiority hypothesis* predicts that the three-
dimensional nature of a live showing (versus two-dimensional photos) of the lineup members, along with having visual information about the full-bodies (not just faces) of the lineup members, would clearly render live lineups superior to photo-lineups. However, a review of the scientific evidence comparing live with photo-lineups (as well as video lineups) showed no support for the live-superiority hypothesis (Fitzgerald, Price, & Valentine, in press). In addition, there are a number of practical difficulties involved in organizing and administering live lineups, including the greater difficulty of finding appropriate fillers for live lineups and the need to carefully orchestrate the timing and roles of various people at the lineup event. Because of the apparent absence of any significant advantage in accuracy along with the greater practical difficulty of live lineups, “live lineups are rarely the best option in practice” (Fitzgerald et al., in press, p. X).

We take no position on live versus photo because preference for one over the other is likely to depend on the circumstances of the particular case. For example, if the witness described something distinctive about the body of the culprit or the culprit’s gait, then maybe a live lineup ought to be preferred. But the difficulty of constructing such a lineup with live fillers who match the witness’s description of the culprit can be very challenging. Also, it is not uncommon for a suspect to be at large, which precludes the use of a live lineup. In other cases, the stress of a victim-witness potentially having to view her assailant live for purposes of identification might cause difficulties that could be avoided with a photo-lineup. Some law enforcement agencies have used live lineups because there were reasons to believe that the witness might be able to identify the culprit’s voice by having lineup members speak which, of course, is not possible while doing a photo-lineup. On the other hand, doing a photo-lineup does not preclude doing a later voice-lineup with that suspect using only recorded voices and no
visual information. In fact, both theory and data converge on the idea that voice identification should be conducted separately from the visual lineup because the diagnosticity of the information obtained is greater if the witness can identify the voice and the face independently of each other (Pryke, Lindsay, Dysart, & DuPuis, 2004). In this important sense, conducting photo-lineups does not preclude the identification of voices using a separate procedure.

**Lineups as Distinguished from Showups**

The 1998 scientific review paper dealt only with lineups. But there is another common identification procedure, called showups, that is included in this new scientific review paper. The basic distinction between lineups and showups is that lineups embed the suspect among known-innocent fillers whereas showups do not use fillers and instead simply present the suspect alone. There is no debate among eyewitness scientists about the fact that lineups produce better outcomes than do showups, whether the outcomes are measured in terms of diagnosticity ratios or measured using signal-detection based methods. There is some debate about the *process* by which the outcomes for lineups are superior to the outcomes for showups. For example, it has been suggested that the use of good lineup fillers helps the witness decide which facial features are relevant for making an identification decision (e.g., see Wetmore et al., 2015; Wixted & Mickes, 2014). Others, however, note that lineups appear to be superior to showups only because a large share of mistaken identifications are siphoned off to fillers when lineups are used whereas *all* mistaken identifications land on the innocent suspect for showups because showups have no fillers (e.g., see Smith, Wells, Lindsay, & Penrod, 2017; Wells, Smith, & Smalarz, 2015). These two accounts of how lineups manage to produce better outcomes than showups are very different, but no eyewitness scientist contests the general observation that a lineup with good fillers results in a better applied outcome.
If there is no question about the superiority of lineups over showups, why do we have recommendations about how showups should be conducted? Why not simply state that showups should never be conducted and that lineups should be conducted instead? There have been calls by some eyewitness scientists for the elimination of showups based on the clear evidence that showups are inferior to lineups (see Levi & Lindsay, 2001). But there are good legal policy reasons to permit showups under certain conditions even though a lineup would be more diagnostic. Showups most commonly occur under conditions in which a person who fits the description of culprit is detained on the street shortly after the occurrence of the crime in an area not far from a crime. But, in the absence of grounds to arrest the detained individual, the individual can only be detained for relatively brief period of time, a period of time that is too short to be able to construct and conduct a lineup. However, there is enough time to do a showup in which the eyewitness is brought to the location where the individual is detained in order to see if the witness will confirm or disconfirm that the detained person is the culprit.

Data clearly indicate that identifications from showups are not as reliable as identifications of suspects from properly constructed and conducted lineups. But in these field circumstances, the choice is not between doing a lineup and doing a showup. If there are no grounds to arrest the detained person, the choice is between a showup and setting the individual free. Setting the individual free could be a significant public safety issue because it might mean that the police are setting a dangerous individual free. Moreover, field data show witnesses who are presented with show-ups actually reject the suspect (“No, that is not the person”) more often than they make a positive identification (“Yes, that is the person”) of the showup suspect (Gonzalez, Ellsworth, & Pembroke, 1993). In this sense, showups can permit a quick exoneration of the detained person with minimal disruption to his or her life.
Hence, although showups are not the preferred method of identification, the relatively low diagnosticity of showups compared to lineups is only one of several factors (e.g., quick exculpation, public safety) that needs to be considered from a legal policy perspective. As we explain in Recommendation 9, however, the dilemma in many instances is not necessarily between doing a showup versus setting the detained person free. After all, there might be other reasons to arrest the person (e.g., illegal possession of a weapon), in which case a proper lineup could be conducted. In addition, as we explain in Recommendation 9, if a showup is conducted, the same kinds of safeguards that apply to lineups (e.g., proper instructions) should be used and the suggestiveness of the procedure needs to be minimized.

As a final note, it should be apparent that there should never be such a thing as a photographic showup. After all, the justification for a showup is that the individual has been detained on the street and there is a very limited time frame for conducting an identification procedure. But if investigators are merely in possession of a photo of a suspect, there is no excuse for not taking the time to embed the photo among filler photos and do a proper photo-lineup.

**Overview of Recommendations**

The 1998 scientific review paper had only four recommendations. The current scientific review paper endorses these same four recommendations but expands the number of recommendations to nine. The original four recommendations in the 1998 scientific review paper were that (1) the identification procedure be *double-blind recommendation* (i.e., the lineup should be administered by someone who does not know which person is the suspect and which are fillers), (2) *pre-lineup instructions* to the witness should emphasize that the culprit might not be in the lineup and that the lineup administrator does not know which person is the possible
suspect, (3) there should be only *one suspect per lineup and the suspect should not stand out* from the fillers based on the witness’s description of the culprit or other factors that would draw attention to the suspect, and (4) a *confidence statement* should be secured from the witness at the time of identification and prior to any opportunity to get feedback about the identification decision. Although we include these four recommendations in our new set of nine, we have tweaked them in certain ways. For example, the double-blind recommendation now includes other means for accomplishing the goal of preventing influence from the lineup administrator that do not necessarily require an additional administrator. The pre-lineup instructions include reworked language that might make them more effective. Also, the securing of a confidence statement includes the idea of taking a confidence statement on rejection decisions instead of taking a confidence statement only if the eyewitness makes an affirmative identification decision.

The original four recommendations in the 1998 scientific review paper were restricted almost exclusively to matters that occurred only during the lineup itself. Some of the six new recommendations, however, cover broader territory. For example, new recommendations concern matters that occur *before* the commencement of an identification procedure, including consideration of when it might be unwise to conduct an identification procedure as well as the problem of repeated identification procedures with the same witness and suspect. In addition, we make a recommendation concerning showups.

The following is a brief description of each of the nine recommendations. The numeric order of the recommendations corresponds roughly to the temporal order in which police/administrators would encounter the matters covered by the recommendation (except for the last recommendation, which concerns showups). The italicized words in brackets at the end
of each description are short names for each recommendation that we will use at certain points in
the current article.
1. There should be evidence-based grounds for reasonable suspicion of an individual prior to
conducting an identification procedure involving that individual. [reasonable-suspicion
recommendation]
2. Prior to conducting an identification procedure, there should be a clear record of the
eyewitness’s description of the culprit, viewing conditions, attention during the crime, and any
claims of prior familiarity with the culprit. [pre-lineup interview recommendation]
3. Lineups should be conducted using a double-blind procedure (i.e., neither the administrator
nor the witness should know who the suspect is in the lineup), or an equally-effective method
of preventing influence from the lineup administrator, so as to prevent inadvertent influence
from the lineup administrator. [double-blind or equivalent recommendation]
4. There should be only one suspect per lineup and the lineup should contain at least five
appropriate fillers who are selected in a way that does not make the suspect stand out in the
lineup based on appearance or background. [lineup fillers recommendation]
5. When inviting an eyewitness to attend a lineup procedure, police should not suggest that a
suspect has been arrested or that the culprit will be present in the identification procedure.
Police should be sure to tell the witness that they do not need to select anyone in the lineup.
The eyewitness should be instructed that the lineup administrator does not know which person
might be the suspect and that the culprit might not be in the lineup at all, so the correct answer
might be “none of these.” Witnesses should also be given the option of “don’t know.”
Witnesses should be told that they will need to state their confidence after making an
identification decision. Witnesses should be told that the investigation will continue even if no 
identification is made. [*pre-lineup instructions recommendation*]

6. If an eyewitness makes an identification, the eyewitness should be asked how confident he or 
she is that the identified person is the culprit. If a witness makes a rejection decision, the 
witness should also be asked how confident he or she is that the culprit is not in the lineup. 
[*confidence statement recommendation*]

7. The entire identification procedure, including pre-lineup instructions and witness 
confidence statement, should be video-recorded. [*video-recording recommendation*]

8. There should be only one identification attempt using the same suspect and the same 
eyewitness no matter reasonable it might seem to conduct a second identification procedure 
with the same suspect. [*avoid repeated identifications recommendation*]

9. Showups should be avoided whenever it is possible to conduct a lineup (e.g., if there are 
grounds to arrest the person independent of eyewitness evidence, then there is time to conduct 
a lineup instead of a showup). In cases for which it is necessary to conduct a showup, it should 
include the same procedural components as a lineup, such as instructing the witness that the 
person to be viewed might not be the culprit, video-recording the entire procedure, and 
securing a confidence statement. [*showups recommendation*]

In the detailed treatments of the nine recommendations that follow, we have attempted to 
place a strong emphasis on communicating an underlying principle for the recommendation. The 
underlying principle for a recommendation is important because there can be times in which 
circumstances might require deviation from the literal specifics of the recommendation, but the 
deviation would still be able to conform to the underlying principle. For example, our 
recommendation for how to select fillers to use in a lineup is relatively specific, but in the end it
is more important that the underlying principle be achieved, namely that the fillers should be
chosen in a way that would not make an innocent suspect stand out in the lineup. The underlying
principle is important for another reason as well, namely the specific recommendation might be
technically followed and yet the principle itself is violated at some other point or level. Consider,
for example, the invitation to view a lineup recommendation; “When inviting an eyewitness to
attend a lineup procedure, police should not suggest that a suspect has been arrested or that the
culprit will be present in the identification procedure.” The underlying principle concerns the
need to avoid communications to the eyewitness that could undermine the pre-lineup instructions
emphasizing that the culprit might not be in the lineup. Although the invitation to view a lineup
might be the most likely time at which this principle is violated, it is not the only possible time.
In this sense, the underlying principle is more important than is the specific wording of the
recommendation.

RECOMMENDATIONS

Each of the nine recommendations begins with a statement of the recommendation. It
then describes the rationale for the recommendation, including relevant data and the logic behind
the recommendation. In addition, most of the recommendations have nuances or caveats, and
some have practical concerns that are discussed.

Recommendation 1: The Reasonable-Suspicion Recommendation

There should be reasonable grounds to suspect that a possible suspect is the culprit prior
to making that person the subject of lineup.

Conducting lineups in the absence of reasonable suspicion (e.g., having only a hunch)
contributes to a low base rate for culprit-present lineups (i.e., a high base rate for culprit-absent
lineups). In the case of lineups, base rate refers to the rate for which the suspect in the lineup is
guilty versus innocent. For nearly 40 years eyewitness researchers have known that culprit-absent lineups are the chief bane of innocent suspects (Wells & Lindsay, 1980). Structurally, a proper lineup contains only one suspect, who might or might not be the culprit (see Recommendation 4). It follows from this structure of lineups that a mistaken identification of an innocent suspect cannot happen with a culprit-present lineup and an identification of the culprit cannot happen with a culprit-absent lineup (Wells & Turtle, 1986). Therefore, low base rates for culprit-present lineups (high base rates for culprit-absent lineups) create fertile ground for mistaken identifications of innocent suspects and reduce the chances of identifying the culprit. Moreover, culprit-absent lineups inflate the rate at which eyewitnesses identify known-innocent fillers, thereby tainting that witness’s credibility for any later lineup that might include the culprit.

The reasonable-suspicion recommendation derives from the observation that there are no laws or other mechanisms in place to prevent jurisdictions from making investigative decisions that result in extremely low base rates for culprit-present lineups (Wells, 2006). In fact, the only study of actual lineups to estimate the base rate for culprit-present lineups in any jurisdiction (in this case the Houston, TX, Police Department) yielded an estimate of a mere 35% (Wixted, Mickes, Dunn, Clark, & Wells, 2016). If this estimate is correct for Houston, then the suspect was innocent in 65% of their lineups. Another field study found that 40% of the lineups (in Northern California jurisdictions) had no pre-lineup evidence at all indicating that the suspect was the culprit; and for an additional 30% of the lineups there was minimal evidence (Behrman & Richards, 2005). Moreover, a national survey of U.S. law enforcement agencies reported that more than one third of the agencies stated that they needed no evidence at all or needed only a mere hunch that a person might be the culprit before placing that person in a lineup (Wise, Safer,
& Maro, 2011). Of course, the base rate is likely to vary from one jurisdiction to the next depending on the practices and policies in place (Wells, 1993). But, at the time of this writing we know of no jurisdiction in the U.S. whose policies or written procedures require, urge, or even mention the idea that there ought to be some form of concrete evidence against the possible suspect before conducting an identification procedure focused on that person.

The Importance of Base Rates

There are many studies in the basic judgment and decision-making literature showing that people struggle to grasp the strong impact that prior probabilities and base rates have on test outcomes (e.g., Kahneman & Tversky, 1973). To illustrate, it is perhaps instructive to draw a close analogy between eyewitness identification testing and medical diagnostic testing. In medical diagnostic testing, it is common for medical organizations to issue guidelines about when to perform diagnostic tests versus forgo such tests. Consider, for example, the prostate-specific antigen (PSA) test for prostate cancer (see Vollmer, 2006). Although the PSA test is just as accurate for men under 30 as it is for men over 50, almost every positive PSA test result on men under 30 is a false alarm whereas only a small fraction of positive PSA test results on men over 50 are false alarms. Because the base rate (or prior probability) that an under-30 male will have prostate cancer is nearly zero, almost every positive result is a false alarm.

This same principle applies to eyewitness identification procedures. In the case of lineups, the base rate is the rate at which the suspect in the lineup is guilty versus innocent. More formal treatments of the (Bayesian) mathematics behind this problem are available (see Wells, Yang, & Smalarz, 2015; Wixted & Wells, 2017), but a simple version of the problem is presented here. Assume that the probability that an innocent suspect will be identified from a culprit-absent lineup is 6% and the probability that a guilty suspect will be identified from a
culprit-present lineup is 60%. Assume as well that the long-term base rate for culprit-present lineups is 50% (and the culprit-absent lineup base rate is therefore 50%). Suppose now that 1,000 lineups were conducted (500 culprit-present and 500 culprit-absent). We would expect 300 identifications of guilty suspects (60% of 500) and 30 identifications of innocent suspects (6% of 500). In this example, 330 suspects are identified and 9.1% of these suspects (30/330 = .091) are innocent.

Now suppose that, instead of a 50% base rate, the base rate were lowered to 30% (300 culprit-present lineups and 700 culprit-absent lineups). Now, the 1,000 lineups would be expected to yield 180 identifications of guilty suspects (60% of 300) and 42 identifications of innocent suspects (6% of 700). The result is that 222 suspects are identified (180 + 42) and 18.9% of these (42/222 = .189) are innocent. In this 30% base-rate example, the percentage of identified suspects who are innocent more than doubles compared to when the base rate is 50%.

Clearly, things get better if the base rate for the suspect being guilty is increased to 70%. At a base rate of 70%, the 1,000 lineups would yield 420 identifications of guilty suspects (60% of 700) and only 18 identifications of innocent suspect (6% of 300). In this example, 438 suspects are identified (420 + 18) and only 4.1% are innocent.

Notice in the above examples that the eyewitnesses themselves are performing just as well when the base rate is 30% as they are when the base rate is 70% (just as the PSA test performs as well when used on men of age 30 as it does on men of age 60). The difference is that the 30% base rate allows for many more false alarms than does the 70% base rate. Every time a culprit-absent lineup is conducted, there is always some probabilistic jeopardy for an innocent suspect. Therefore, minimizing the chances of presenting witnesses with culprit-absent lineups is one way to reduce the problem of wrongful convictions.
Even when the witness does not identify the innocent suspect in a culprit-absent lineup, culprit-absent lineups strongly increase the tendency of eyewitnesses to identify a known-innocent lineup filler (Clark & Wells, 2008; Wells & Lindsay, 1980; Wells & Olson, 2002; Wells, Yang, & Smalarz, 2005). Most filler identifications are made with low confidence, clearly signaling their error-prone nature. However, some are made with higher confidence. Having already confidently identified a known-innocent person as the culprit, it can be argued that the eyewitness is tainted for purposes of any later identification procedure (Wells, Steblay, & Dysart, 2012). Having already confidently identified a known-innocent person as the culprit, it can be argued that the eyewitness is tainted for purposes of any later identification procedure (Wells, Steblay, & Dysart, 2012). Suppose, for example, an eyewitness is shown a culprit-absent lineup and confidently identifies a filler. Later, police receive information about who the actual culprit is. It is too late to undo the fact that the witness has already confidently identified someone else, a known-innocent person, as the culprit. This prior identification of an innocent person makes the prosecution of the newly identified person substantially more difficult for prosecutors.

What is Reasonable Suspicion?

Science cannot define what reasonable suspicion ought to be any more than science can (or should) define probable cause or other legal concepts. Nevertheless, our use of the term reasonable suspicion is intended to mean something different than probable cause. Probable cause is a higher standard that implies that there is enough evidence to arrest an individual. Instead, by “reasonable suspicion” we mean that there should be some tangible form of evidence that can be articulated above and beyond a mere hunch. It has been suggested that a reasonable-suspicion requirement could be easily implemented by requiring detectives to present their case for conducting a lineup to a chief of detectives, who could question the detective about why this
lineup is being conducted with this particular person as its focus (Wells, 2006). Ideally, the detective would be able to point to something concrete; if not, the chief of detectives might suggest that the detective investigate further in order to have more confidence that the witness is not going to be shown a culprit-absent lineup.

**Final Comments on Reasonable Suspicion**

The medical field’s understanding of import of the impact of base rates on medical diagnostic test outcomes is far ahead of the legal system’s understanding of the impact of base rates in eyewitness lineup test outcomes. Furthermore, the concern about base rates in eyewitness identification might be even more important than in diagnostic medical tests because diagnostic medical tests can be repeated with that patient to confirm reliability of the result, or a different type of test can be performed to look for convergence of results. An eyewitness identification test cannot be repeated with that same witness and same suspect (see Recommendation 8); so it is important to ensure that the chances of presenting eyewitnesses with a culprit-absent lineup are not unduly high.

**Recommendation 2: Pre-lineup Interview Recommendation**

*Immediately following the event in question, investigators should interview the witness to obtain an accurate and complete description of both the event and the culprit. In doing so, investigators should also document the witness’s viewing conditions and attention during the crime, any claims of prior familiarity with the culprit, whether someone suggested to them who the culprit is, whether they have interacted with other witnesses regarding the identity of the culprit, and whether they have seen any news accounts of the event. When cases involve multiple witnesses, investigators should instruct witnesses not to discuss their accounts with one another.*
Recommendation 2 relates to the conduct of an initial interview with a witness or victim, during which time an investigator collects an initial statement relating to the person’s memory for the event and the culprit(s). Collection of a detailed description of the culprit is a critical form of evidence that can facilitate investigators’ attempts to locate a suspect (Brown, Lloyd-Jones, & Robinson, 2008; Kebbell & Milne, 1998). There is now a substantial research literature on the most effective procedures for interviewing a witness or victim following an event (see Dando, Geiselman, MacLeod, & Griffiths, 2015; Fisher, Schreiber-Compo, Rivard, & Hirn, 2014), as well as the influence of suggestive or misleading interviewing procedures that should be avoided (see Brainerd & Reyna, 2005; Loftus, 2017; Newman & Garry, 2013). Specific interviewing procedures have also been developed for the collection of *person descriptions* (see Demarchi & Py, 2009; Gabbert & Brown, 2015; Meissner, Sporer, & Schooler, 2009; Sporer, 1996).

Although the current review focuses largely on the adult witness literature, many of the same principles of memory apply to the interviewing of child witnesses (who are particularly susceptible to suggestion). A robust literature is available for interested readers documenting the challenges of interviewing child witnesses (see Kask & Bull, 2009; LaRooy, Heydon, Korkman, & Myklebust, 2015), including the development of specific, evidence-based approaches for interviewing children (see LaRooy, Brubacher, et al., 2015; Sternberg, Lamb, Esplin, Orbach, & Hershkowitz, 2002).

**Contents and Accuracy of Person Descriptions**

Archival studies suggest that witnesses tend to provide between 7 and 9 descriptors of a perpetrator on average, frequently including information about perceived height, weight, gender, ethnicity, and age (Fahsing, Ask, & Granhag, 2004; Granhag, Ask, Rebeilius, Ohman, & MacGiola, 2012; Sporer, 1992; van Koppen & Lochun, 1997; Yuille & Cutshall, 1986).
Descriptions of the perpetrator’s clothing, stature, and facial features are generally less frequent. When specific facial descriptors are provided, the majority refer to upper regions of the face, in particular the hair, eyes, and nose. Although estimates of height, weight, and age can be biased by the witness’s own characteristics (e.g., individuals who are less than average height tend to underestimate height; cf., Flin & Shepherd, 1986), witnesses otherwise appear to provide an accurate, general impression of the perpetrator. Such descriptions, however, are often lacking in specific details (Douglass, Brewer, Semmler, Bustamante, & Hiley, 2013; Fahsing et al., 2004) that might prove useful for the construction and assessment of identification arrays (Corey, Malpass, & McQuiston, 1999), and it is therefore important that investigators use evidence-based procedures to enhance the quality of witnesses’ accounts.

Factors that Influence the Quality of Witness Accounts

Much like other memory phenomena, a host of factors can influence the accuracy and completeness of witness’s memory for the event and perpetrator (see Meissner et al., 2009; Granhag, Ask, & MacGiola, 2013). Consistent with the general eyewitness literature, a distinction can be drawn between system and estimator variables (Wells, 1978). With respect to the former, witnesses appear quite susceptible to the misinformation effect (see Loftus, 2017; Newman & Garry, 2013) in which leading or suggestive questioning from an investigator can distort subsequent memory reports. Exposure to media coverage of an event prior to an interview can also lead witnesses to recall incorrect details that were suggested or inferred (Crombag, Wagenaar, & van Koppen, 1996). Witnesses can be induced to self-generate errors in their descriptions when forced or encouraged to provide a “complete” description of the event or perpetrator (Ackil & Zaragoza, 1998; Meissner, Brigham, & Kelley, 2001), and the use of facial feature checklists also subtly force responses that produce less accurate person descriptions.
(Wogalter, 1991; 1996). Finally, exposure to inaccurate information from other witnesses can also introduce systematic errors in memory (Gabbert, Memon, & Allan, 2003; Gabbert, Memon, Allan & Wright, 2004) and enhance the suggestive nature of investigator questions (Shaw, Garven, & Wood, 1997). Investigators are strongly encouraged to avoid suggestive or leading interviewing practices and to encourage witnesses not to discuss their accounts with one another.

With respect to naturally occurring estimator variables, factors at the time of encoding such as illumination, distance, and time of exposure can influence the quality of person descriptions (Sporer, 1992; van Koppen & Lochun, 1997; Yarmey, 1986; Yarmey, Jacob, & Porter, 2002). Factors that influence a witness’s focused attention also affect witness recall, including their intent to encode the perpetrator (Yarmey, 2004) and the presence of a weapon that draws attention away from the perpetrator’s appearance (Fahsing et al., 2004; Pickel, 1998, 1999). The consumption of alcohol or drugs by a witness can reduce the amount of information provided (Read, Yuille, & Tollestrup, 1992; Schreiber Compo et al., 2017; Yuille & Tollestrup, 1990). Extensive delays between encoding and the time of interviewing can also influence the amount of detail provided (Ellis, Shepherd, & Davies, 1980; Meissner, 2002; van Koppen & Lochun, 1997). Prior familiarity with the perpetrator (i.e., an individual known to the witness) also influences the description (and identification) process (Bruce, 2012). Given the influence of these estimator variables on both the quality of person descriptions and subsequent attempts to identify the perpetrator from a lineup, it is recommended that investigators clearly document the presence of these factors in their report.

Methods for Interviewing Witnesses and Victims

Acquiring a complete, yet accurate, statement from the witness is critical to furthering an investigation. Considerable research has now documented the most effective methods for
interviewing a witness or victim (Dando et al., 2015; Fisher et al., 2014). In general, it is common for investigators to invite an open-ended response from the witness, followed by specific probes associated with key details such as the perpetrator’s physical characteristics (e.g., height, build, age, race, sex, etc.), clothing, or any distinguishing characteristics (Brown et al., 2008; Launay & Py, 2015; Wise, Safer, & Maro, 2011). The use of open-ended, non-suggestive questioning tactics (Clarke, Milne, & Bull, 2011; Walsh & Bull, 2010b) is highly recommended for eliciting a complete narrative from the witness. Although the use of specific probes can increase the number of details provided, the details may come at the expense of lower accuracy of responding (Sauerland, Krix, van Kan, Glunz, & Sak, 2014). As such, caution should be used when moving to closed-ended or two-alternative questions, and (as noted above) the use of suggestive/leading prompts should be avoided altogether.

There are specific interviewing protocols that both avoid the pitfalls of leading and suggestive questioning and enhance reporting by facilitating the retrieval of information from memory. One of the most notable and empirically validated protocols is the Cognitive Interview (CI; Fisher & Geiselman, 1992). A robust literature has demonstrated the effectiveness of the CI for eliciting both detailed event narratives and person descriptions from cooperative witnesses (Memon, Meissner, & Fraser, 2010). Several instructional and mnemonic aspects of the CI appear to be particularly useful: (a) encouraging witnesses to “report all” of the information they can recall but not to guess about anything they are unsure of (e.g., Clifford & George, 1996), and (b) using context reinstatement by asking witnesses to close their eyes and think back to the event context (e.g., Smith-Spark, Bartimusa, & Wilcock, 2017; Vredeveldt, Baddeley, & Hitch, 2012, 2014). We strongly encourage investigators to use the CI protocol when interviewing witnesses and victims.
Another specific protocol for eliciting person descriptions, termed the Person Description Interview (PDI), significantly increases the quantity of person descriptors provided by witnesses (Demarchi & Py, 2009; Demarchi, Py, Groud-Than, Parain, & Brunel, 2013). The PDI incorporates two key instructions to the witness with respect to describing the person of interest: (a) to provide general information about the person before moving to specific featural aspects of the face, and (b) when describing the face to begin with the lower portions (chin and lips) and to move up to the top portions (eyes and hair). Consistent with PDI instructions, encouraging witnesses to provide more general, coarse-grained information during an interview can enhance the quantity of information absent a cost to accuracy (Brewer, Vagadia, Hope, & Gabbert, 2018).

Finally, although witness descriptions are frequently collected via an oral interview conducted by the investigator, at times a witness may be asked to directly provide a written statement. There is mixed evidence with respect to how the format for eliciting a witness’s recall might influence the quantity and quality of information provided, with some studies suggesting that oral interviews produce more information from witnesses (Kraus, Zeier, Wagner, Paelecke, & Hewig, 2017; Sauerland & Sporer, 2011) and others finding no difference as a function of modality (McPhee, Paterson, & Kemp, 2014; Sauerland et al., 2014). A recently developed self-administered interview protocol, referred to as the Self-Administered Interview (SAI; see Hope, Gabbert, & Fisher, 2011) allows cooperative witnesses to self-generate high-quality descriptions of their experience. The SAI prompts witnesses to recall details of the event, including a request for a detailed person description (e.g., hair, complexion, build, distinguishing features, etc.). To facilitate recall, the SAI incorporates the two key elements of the CI noted above – a report everything instruction and a context reinstatement prompt. The SAI significantly increases the quantity of person descriptors when compared with a standard free recall prompt, at a level
comparable to that of the CI (Gabbert, Hope, & Fisher, 2009; Hope, Gabbert, Fisher, & Jamieson, 2014).

**Does Eliciting a Description of the Perpetrator Influence a Witness’s Subsequent Identification?**

For several decades, researchers have explored the influence of generating a person description on subsequent identification of a face. Early research on the verbal overshadowing effect (Schooler & Engstler-Schooler, 1990) suggested that generating a description could subsequently *impair* accuracy on a lineup identification task (when compared with those who did not generate a description). Although the replicability of this effect has been questioned, both a meta-analysis of the literature (Meissner & Brigham, 2001b) and a large-scale replication across multiple laboratories (Alogna et al., 2014) have confirmed the verbal overshadowing effect on accuracy in target-present lineups. Studies have also demonstrated that the effect leads to less discriminability when target-absent lineups are included (Meissner, 2002; Smith & Flowe, 2015; Wilson, Seale-Carlisle, & Mickes, 2018). The verbal overshadowing effect appears to be most robust when witnesses are asked to provide a very detailed, elaborate description of the face (MacLin, Tapscott, & Malpass, 2002; Meissner, 2002; Meissner et al., 2001; Smith & Flowe, 2015). Delays in providing a description of the target person following an event (as little as 20 min) also appear to increase the likelihood of the effect, likely as a result of degraded memory influencing the quality of the witness description (see Alogna et al., 2014; Meissner, 2002; Meissner, Sporer, & Susa, 2008; Wilson et al., 2018). Repeated requests for a description (immediately following an event and immediately prior to an identification attempt) also impair identification performance (Meissner, 2002). Investigators should attempt to collect a witness
description of the culprit as soon as possible following the event using interview procedures that avoid the use of forced or suggestive questioning.

In contrast to the verbal overshadowing effect, other studies have demonstrated that recognition of faces can be facilitated (or enhanced) by asking people to provide a verbal description prior to test (Brown & Lloyd-Jones, 2005; 2006). One factor that can distinguish the likelihood of verbal overshadowing versus facilitation within an eyewitness context involves asking participants to engage in featural versus holistic processing prior to identification. For example, Macrae and Lewis (2002) manipulated processing via a Navon figures task and found that a global, holistic processing perspective facilitated subsequent lineup identification, whereas a local, featural processing perspective inhibited performance (see also Darling, Martin, Hellman & Memon, 2009; Finger, 2002; Perfect, 2003; Perfect, Dennis & Snell, 2007). Recent studies that asked people to describe faces holistically have similarly subsequently demonstrated enhanced identification performance (Wickham & Lander, 2008) and increased quality of composite face production (Frowd, Bruce, Smith, & Hancock, 2008). The potential benefits of holistic interviewing for person descriptions on later identification should be considered (see Gabbert & Brown, 2015).

**Recommendation 3: Double-blind (or Equivalent) Recommendation**

*Lineups should be conducted using a double-blind procedure (i.e., neither the administrator nor the witness should know who the suspect is in the lineup) or an equally-effective method of preventing inadvertent influence from the lineup administrator.*

A lineup administration is a social interaction between a witness and an administrator. Like in other social situations, interpersonal expectancies operate in the context of a lineup administration. Why is the social interaction aspect of lineup administration concerning? When
someone has an expectation about how another is likely to behave, this expectation can cause the person with the expectation to behave differently toward the target of the expectation. This change in the expectation-holder’s behavior elicits the very behavior that was expected from the target (Harris & Rosenthal, 1985; Rosenthal, 2002; Snyder & Swann, 1978). The social interaction that takes place during the administration of a lineup administration is not immune from this interpersonal expectancy phenomenon. There is no presumption that the influence of the lineup administrator is intentional or even that the lineup administrator is aware of the influence.

In essence, a lineup is a test of the hypothesis that a suspect who has been identified by the police is in fact the perpetrator of the crime viewed by a witness. The lineup administrator is fundamentally an experimenter who is conducting a procedure to test this hypothesis. Because people have a tendency to test hypotheses in a way that will confirm their expectations (e.g., Klayman & Haw, 1987; Skov & Sherman, 1986), lineup administrators, like any other experimenters, should follow protocols that will prevent them and their expectations from influencing the results of their tests. Double-blind testing, in which the lineup administrator does not know which person is the suspect and which are merely fillers (e.g., a blind administrator), is the best way of ensuring that any information that administrators have about which lineup member is the suspect will not influence the witnesses’ behavior, including any identification decision they might make or their confidence in that decision. Double-blind testing can also prevent administrators’ expectations from influencing their reports of witnesses’ behaviors during the procedure. In contrast, single-blind lineup administration, in which the administrator knows which lineup member is the suspect and which are fillers (i.e., non-blind administrator),
allows for the possibility that the administrator will communicate to the witness through intentional or unintentional verbal or non-verbal behaviors who the suspect is.

Although the double-blind recommendation is primarily focused on keeping knowledge about the suspect from the administrator of a lineup so that it cannot influence the administrator’s behavior while conducting the identification procedure, it is equally important to keep information about the suspect from the witness (hence the recommendation for double-blind administration). It might seem odd to explicitly warn against letting a witness know who the suspect is before they make an identification decision. In practice, however, we find it not uncommon for circumstances surrounding the identification procedure to alert the witness to which lineup member is the suspect. For example, the witness may be tipped off to who the suspect is after being presented with multiple photo arrays that share only one lineup member (the suspect) in common, which is one of the many reasons for our recommendation to avoid repeated lineup procedures (Recommendation 8).

This recommendation for double-blind administration of lineups was included among the four recommendations made in the previous scientific review paper (Wells et al., 1998). At the time that paper was written, however, there were no studies that directly tested whether a lineup administrator’s knowledge of which lineup member was the suspect influenced witness identifications. Without studies directly testing the effects of double-blind administration of lineups, the recommendation was made based on generalizations from basic studies on experimenter expectancy effects (Harris & Rosenthal, 1985) and early research suggesting that positive feedback to witnesses after they choose the suspect increases their confidence in that choice (Wells & Bradfield, 1998). Since the previous scientific review paper was written, a number of studies have been conducted demonstrating that the single-blind administration of
lineups increases the likelihood that witnesses will identify the suspect (for a review, see Kovera & Evelo, 2017), irrespective of whether the suspect is the culprit (Charman & Quiroz, 2016; Greathouse & Kovera, 2009) or an innocent suspect (Charman & Quiroz, 2016, Greathouse & Kovera, 2009; Zimmerman, Chorn, Rhead, Evelo, & Kovera, 2017).

Scholars have developed several paradigms to examine the effects of administrator influence on witness decisions. In one paradigm, which has been termed the steering paradigm (Kovera & Evelo, 2017), the lineup administrator is a confederate of the experimenter who intentionally engages in behaviors that steer that witness toward the suspect (e.g., Rhead, Rodriguez, Korobeynikov, Yip, & Kovera, 2015) or encourage the witness to make an identification (Clark, Brower, Rosenthal, Hicks, & Moreland, 2013). However, the influence of administrators on witnesses need not be intentional. In the cue-disruption paradigm, all administrators know who the suspect is but half of them are prevented from sending cues (whether intentionally or unintentionally) to the witness during the administration of the lineup. In one study, for example, the contact between the administrator and the witness was limited by having the administrators stand behind the witnesses while they viewed a photo-array. Witnesses were less likely to identify the suspect when the administrator stood behind them than when the administrator sat in front of or beside the witness (Haw & Fisher, 2004). In the double-blind paradigm, participants are randomly assigned to be either witnesses or lineup administrators; half of the lineup administrators are told who the suspect is and the other half are not (Phillips, McAuliff, Kovera, & Cutler, 1999). These participant administrators then present the photo-array to the participant witnesses. Across all paradigms, when administrators know who the suspect is and are not prevented from sending cues to the witness, witnesses are more likely to choose the suspect from the lineup, whether the suspect is the culprit or not (Kovera & Evelo, 2017).
These studies tell us that changes in the behavior of administrators during the administration of the lineup are responsible for this increase in witness picks of the suspect. In the steering paradigm studies, the administrators’ behaviors were intentionally manipulated to steer the witness toward the suspect and away from fillers (e.g., Rhead et al., 2015). In double-blind paradigm studies, observers reported that non-blind administrators placed more pressure on witnesses to choose someone from the lineup than did blind administrators (Greathouse & Kovera, 2009) and that pressure was directed toward choosing the suspect rather than a filler (Zimmerman et al., 2017). Non-blind administrators were more likely to directly ask witnesses about the suspect than were non-blind administrators (Zimmerman et al., 2017). Non-blind administrator influence can be non-verbal as well; non-blind administrators are also more likely than blind administrators to smile when a witness is looking at the suspect rather than a filler (Charman & Quiroz, 2016; Zimmerman et al., 2017).

These differences in behaviors between blind and non-blind administrators affect which photo witnesses choose from lineups, not whether they make a choice at all. Witnesses are equally likely to choose someone from a lineup, irrespective of whether the lineup administrator knows who the suspect is (Greathouse & Kovera, 2009; Kovera & Evelo, 2017). The increase in witness identifications of the suspect from single-blind lineup administrations appears to be the result of witnesses who would have identified a filler (and do so under blind administration) identifying the suspect instead due to influence from the non-blind administrator (Kovera & Evelo, 2017). This pattern of findings, replicated in a number of studies (Charman & Quiroz, 2016; Greathouse & Kovera, 2009; Kovera & Evelo, 2017), is known as the filler-to-suspect shift and provides compelling evidence that single-blind lineup administration allows
administrators to transmit information about who the suspect is to witnesses, even if unintentionally.

In addition to affecting witnesses’ identification decisions, single-blind lineup administration allows administrators to provide feedback to witnesses about their decisions. Non-blind administrators react to witness identifications in ways that send information to witnesses about whether their choice was “correct” (i.e., an identification of the suspect; Charman & Quiroz, 2016; Garrioch & Brimacombe, 2001). Two decades of research supports the conclusion that providing feedback to witnesses that they identified the suspect increases their confidence in the accuracy of their decision, especially among eyewitnesses who have made a mistaken identification (Steblay, Wells, & Douglass, 2014; see Recommendation 6 for a more complete discussion of this research). This confirming feedback effect attenuates the relationship between confidence and accuracy (Bradfield, Wells, & Olson, 2002), rendering witnesses’ reports of their confidence useless for judging their accuracy (Wixted & Wells, 2017). In addition to preventing administrators from providing feedback that will influence witnesses’ reports of their confidence, double-blind administration will also prevent other unwelcome effects of feedback such as the contamination of witnesses’ memory for the conditions under which they witnessed the crime (Steblay et al., 2014), the impairment of witness memory for the perpetrator (Smalarz & Wells, 2014a), and lessening the ability of jurors to differentiate between accurate and inaccurate witnesses (Smalarz & Wells, 2014b).

Knowing who the suspect is may also influence what information administrators record about witnesses’ behavior during the identification procedure. Even though we recommend video-recording the lineup administration (see Recommendation 7), currently most lineups in the U.S. (> 75 percent) are not video-recorded (Police Executive Research Forum, 2013). In these
non-video-recorded cases, the only contemporaneous record of what happened during the procedure is information recorded by the administrator. Often, lineup administrators will not remake a record of the verbatim statement of the eyewitness but instead make a note of the “gist” of what the eyewitness said. It is possible that knowing who the suspect is may change how administrators assess and record witnesses’ choices from lineups. If so, when witnesses make tentative identifications (e.g., “I don’t know. I think it may be number 4, but I’m not certain.”), for example, administrators who know that the witness is talking about a suspect may record a positive identification of the suspect but administrators who know that the witness is talking about a filler may record the very same behavior as a non-identification (Rodriguez & Berry, 2014).

Although there are limited empirical data that directly bear on the effects of administrators’ knowledge of who the suspect is on their reports, evidence continues to mount that forensic examiners’ expectancies influence their evaluations (for a review, see Kassin, Dror, & Kukucka, 2013). In addition, there are data from both laboratory and field studies suggesting that administrators who know who the suspect is record witness choices differently than do administrators who lack this knowledge. Indeed, for five years in Queens County, NY, the District Attorneys’ Office recorded choices made by witnesses from live single-blind lineups. Choices were only recorded when the administrators judged that it had been made with a high degree of confidence and was not tentative (as reported in Mecklenburg, 2006). This procedure resulted in a very low rate of filler identifications (between .56% and 5.62%) in comparison to the rates of filler identifications from double-blind lineups in other field studies, which more than doubled these rates (Klobuchar, Steblay, & Caligiuri, 2006; Wells, Steblay, & Dysart, 2015). In other field data from the Evanston Police Department in Illinois, reports from double-blind
sequential lineups were more likely to be verbatim reports of witness statements than were the reports after single-blind sequential lineups (83% vs. 39%, Steblay 2011). The interpretation of the Evanston data is problematic because only double-blind administrators were instructed to record what words the witnesses used to make their identification and single-blind administrators were not. Only one experiment has tested how administrators make records of the behavior of eyewitnesses as a function of whether the administrator of a lineup was blind and whether a confederate-witness chose the suspect or a filler (Rodriguez & Berry, 2014). Although double-blind administrators were just as likely to report that witnesses had made a positive identification when the witness identified a filler as they were when the witness identified the suspect, single-blind administrators were more likely to report incorrectly that witnesses who identified a filler had not made an identification.

These effects of single-blind lineup administration witnesses’ identification decisions, their confidence, and administrators’ reporting behavior support the use of double-blind procedures when collecting eyewitness identification evidence. The Department of Justice (NIJ, 2017) and the National Academy of Sciences (2014) have made similar recommendations. Nevertheless, many jurisdictions have yet to put this procedure, or a similar alternative, into practice (Kovera & Evelo, 2017).

Two alternatives to the use of blind administrators have been identified in state statutes: automated computer administration and the folder-shuffle method. However, only the automated computer method (not the folder shuffle method) fully controls the problems that double-blind lineup administration controls (Kovera & Evelo, 2017). An automated computer program removes the police officer from the lineup procedure. If it is designed appropriately, it will not provide cues to witnesses about which lineup member is the suspect. Using a computer program
not only ensures that lineups are presented to witnesses uniformly, the programs can also record many aspects of the administration, including the photos shown, the witnesses’ choices, their confidence, and the time it took for them to make an identification decision. In addition, the computer can automatically randomize the order of the photos and record the location of the photos in the case of simultaneous presentation or the order in which the photos were shown in the case of sequential presentation.

In the folder-shuffle method, administrators number folders containing the photographs and then shuffle their order so that they do not know which folder contains the suspect. The witness then looks at the photos in the folders, either one at a time (sequentially) or at the same time (simultaneously). The administrator can leave the room, after providing appropriate instruction, while the witness looks at the pictures. At a minimum, the non-blind administrator should stand behind the witness in a position that prevents viewing the photos. Although the folder-shuffle method has been recommended by some, it does not completely eliminate the opportunity for non-blind administrators to influence witnesses’ confidence or reports of their witnessing conditions if they have contact before these reports are made. Thus, the folder-shuffle method may not completely eliminate administrator influence. Therefore, we recommend that the police make use of an administrator who does not know who the suspect is, even if that means calling on neighboring police departments to provide the blind administrator in cases in which everyone in a given department has knowledge of a particular case. New Jersey, for example, has mandated double-blind administration since 2002 and has been able to implement successfully the mandate through the use of cooperative agreements to loan officers to nearby departments for the purpose of conducting blind identification procedures despite having departments with as few as two officers. To the extent that their experience is representative of
what police departments in other states would experience, it should be possible to implement the double-blind procedure even in small jurisdictions.

In addition to concerns about limited resources making it difficult to implement double-blind procedures, some have objected to their adoption because of the loss of correct identifications associated with blind administration (Clark, 2012a; 2012b). Although it is true that double-blind procedures reduce both correct and mistaken identifications, they do so by eliminating the opportunity for administrators to cue the witness to which lineup member is the suspect. Given that the legal system requires that an eyewitness identification be based on the independent memory of the witness (*Perry v. New Hampshire*, 2012), a loss of an identification obtained through administrator influence should not be a concern. Indeed, some scholars have termed these correct identifications obtained through suggestive procedures to be illegitimate hits (Wells, Steblay, & Dysart, 2012). Thus, double-blind procedures serve to protect suspects’ rights to due process.

**Recommendation 4, Lineup Fillers Recommendation**

*There should be only one suspect per lineup and the lineup should contain at least five appropriate fillers who are selected in a way that does not make the suspect stand out in the lineup based on the lineup members’ physical appearances, or other factors such as clothing or background.*

Recommendation 4 concerns what might be considered the most widely known problem that can afflict lineups, namely a lineup that is constructed in a way that makes it obvious which member is the suspect. This idea is the source of scores of cartoons and jokes about the perceived failings of criminal justice, such as one that depicts a person embedded in a lineup composed of a
dog, cat, refrigerator, and a microwave oven. Nevertheless, jocular treatments of biased lineups hide a serious problem and this problem has much greater complexity than meets the eye.

The problem of the biased lineup is one of the oldest in the scientific study of eyewitness identification. In fact, the first published experiment on eyewitness identification that manipulated the presence versus absence of the culprit in the lineup (a now routine feature of eyewitness identification experiments) was an experiment that also manipulated the lineup fillers to be either similar or dissimilar to the suspect (Lindsay & Wells, 1980). Not surprisingly, in culprit-absent lineup conditions the use of high-similarity lineup fillers strongly reduced mistaken identifications of the innocent suspect compared to the use of low-similarity fillers. In culprit-present conditions, the use of these same high-similarity fillers had only a minor impact on accurate identifications of the culprit relative to the use of low-similarity fillers. This pattern of results, showing that using low-similarity fillers increases the chances of mistaken identification of an innocent suspect, has been repeatedly replicated.

Despite the relative ease of replicating the basic finding that low-similarity fillers increase the risk of mistaken identification of an innocent suspect, there is not total agreement among eyewitness scientists regarding the best strategy for choosing fillers to serve in a lineup. There are two primary strategies for selecting fillers (Luus & Wells, 1991). One strategy uses the verbal description that the eyewitness gave of the culprit (e.g., “white male, mid 20s in age, about 5 feet 10 inches tall, short dark hair, no facial hair, medium build”). This method of choosing fillers was called the match-to-description strategy. The other strategy, called the resemble-suspect strategy, involves choosing fillers who physically resemble the suspect. The resemble-suspect strategy can be problematic because it has no criterion or “stopping point” for deciding how similar the fillers should be, resulting in lineup fillers who are too similar, causing
a different problem than a biased lineup (Luus & Wells, 1991). In effect, extremely high similarity creates a lineup of near-clones, thereby making it too difficult to identify the culprit from a culprit-present lineup. The match-to-description strategy, in contrast, has a natural stopping point (the description) and does not risk creating such high levels of similarity between fillers and the suspect that it would interfere with obtaining accurate identifications of the culprit (Luus & Wells, 1991). An experiment comparing the two strategies to a biased (low similarity) lineup showed the two strategies to be equally effective in reducing innocent suspect identifications, but the resemble-suspect strategy produced a reduction in accurate identifications of the culprit whereas the match-description strategy did not. However, other studies showed no detrimental effect on culprit identifications from using the resemble-suspect strategy and no evidence that it made the innocent suspect stand out (Timmeliff & Clark, 2000) or no advantage one way or the other for match-description versus suspect resemblance strategies (Darling, Valentine, & Memon, 2008).

In a meta-analysis of the data on filler similarity, lineups classified as high similarity produced a reduction in culprit identifications relative to low-similarity lineups but not relative to moderate similarity lineups (Fitzgerald, Price, Oriet, & Charman, 2013). As noted in the meta-analysis, the categorization of low, medium, and high similarity lineups reflected relative rather than absolute levels of similarity—the concept of similarity cannot be defined in absolute terms, with clear criteria for each level of sameness, rather its meaning is grounded in relative terms. In a more recent study, fillers who resembled the suspect were selected from either an extremely large database of faces (which produced very high similarity fillers) or a more modest sized database (which produced more moderate levels of similarity; Bergold & Heaton, 2018). Compared to the more modest size database of faces, using the large database of faces for
selecting fillers resulted in a reduction in accurate identifications of the culprit by producing too much similarity between the fillers and the suspect. Overall, the data suggest that using the resemble-suspect strategy could produce too much similarity between the suspect and the fillers such that it interferes with identifications of the culprit, especially when fillers are selected from large databases. This problematic condition could become more prevalent as the ability to rapidly search large databases of faces becomes increasingly more common.

A Blended Approach for Choosing Fillers

The net result of these complex problems is that the science has not yet been able to specify what the optimal level of similarity (of fillers to the suspect) ought to be and thus, at this time, there is no single strategy or formula for choosing fillers to be used in a lineup. Nevertheless, there are generally accepted broad ideas about how to choose fillers for a lineup and they tend to involve a blend of the match-description and resemble-suspect strategies. First, there is general agreement among experts that a minimal requirement is that the fillers fit the description that the eyewitness gave of the culprit. A failure to match the witness’s description might introduce a serious bias even though the members may look very similar. For example, a witness might describe the culprit as “male, 40s, shaven head, striking blue eyes, very solid build around the neck and shoulders.” A subsequent lineup could have an array of people who are very similar looking but if only one had the blue eyes, that lineup member becomes a real standout.

An exception to the match-description needs to be made when the suspect him or herself does not fit that description. For example, a person might become a suspect for reasons other than their appearance. Furthermore, if the description of the culprit mentioned a moustache but the suspect does not have a moustache, then the fillers also should not have a moustache. In other words, if there is a discrepancy on some physical feature between the eyewitness’s description of
the culprit and the appearance of the suspect, the fillers should match the suspect’s appearance (rather than the witness’s description of the perpetrator) on that feature.

Sometimes, a suspect has a unique feature, such as a tattoo or a scar. There are two general approaches to dealing with this problem, made easier by advances in technology. One is to duplicate this feature on the fillers, which could be done electronically in the case of photolineups. The other approach is to cover the unique feature on the suspect and then place that same-looking cover on each of the fillers at the same location on their bodies. Both the “duplicate” and “cover” approaches appear to be equally effective (Colloff, Wade, & Strange, 2016).

How much similarity between the fillers and the suspect is too much similarity? Unfortunately, the research data do not provide an answer to that question, but a reasonable rule of thumb might be that the fillers should not be so similar to the suspect that someone who knows the suspect would struggle to pick the suspect out of the lineup.

Another complication in selecting fillers to be used in lineups can occur when a person becomes a suspect based on resemblance to a face composite (e.g., forensic sketch or computer-generated face) or a surveillance image. If the person became a suspect based on resemblance to the composite or surveillance image whereas the fillers were chosen based merely on their match to the witness’s verbal description, then there is a risk that the suspect will stand out. Hence, in such cases, fillers for a lineup need to be chosen based on their similarity to that same composite or surveillance image rather than chosen based on the verbal description given by the eyewitness (see Wixted and Wells, 2017).

There is broad agreement that the lineup that is created should be able to pass a “mock witness” test. A mock witness test is one in which a large number of people are individually
given the description that the witness had given of the culprit, shown the lineup, and asked which person they think is the suspect. The ideal outcome from such a test would be if the suspect were picked by these mock witnesses only 1/Nth of the time, where N is the nominal number of lineup members. So, for a six-person lineup, a good outcome would be if the suspect were picked 1/6th of the time. Consider again the witness description “male, 40s, shaven head, striking blue eyes, very solid build around neck and shoulders” and the suspect is the only one with blue eyes. This is likely to result in most mock witnesses choosing the suspect rather than spreading their choices across the lineup members. We are not suggesting that police have to conduct a mock witness test on each lineup in order to know if they have a good lineup. Instead, we believe that a conscientious and objective detective would have a good sense of whether the lineup was fair without conducting a mock witness test with a large number of people. However, we recommend that a non-blind police officer building the lineup have at least one or two other people (ideally, blind as to which person is the suspect) look at the witness description and the lineup to get a second opinion on whether it would pass a mock witness test.

Mock witness tests have been around since the 1970s (Doob & Kirshenbaum, 1973) and a number of different statistics have been developed to estimate lineup bias from mock witness results (e.g., see Wells, Leippe, & Ostrom, 1979; Malpass, 1981; Tredoux, 1998). Mock witness measures tend to predict choices of eyewitnesses from culprit-absent lineups (Tredoux, Parker, & Nunez, 2007), but mock witness tests have limits (see Wells & Bradfield, 1999). One of these limits is that a mock witness test is insensitive to whether the level of similarity between fillers and the suspect is too high, so high that it would strongly harm rates of accurate identifications if the suspect is the actual culprit. For example, a lineup of clones would produce a good result.
from a mock witness test (1/6th of choices are of the suspect), even though it would not be a good lineup.

**Background, Clothing and other Factors**

The physical characteristics of the fillers are not the only factors that can make the suspect stand out in a lineup. In photo-lineups, for example, the background of the photos, the size or brightness of the images, and the source of the photo could make the suspect’s photo stand out from the others. Backgrounds are relatively easy to fix on photos with simple modern editing software. Similarly, clothing can easily be eliminated from headshots with editing software. Sometimes, a suspect’s photo is from a different source (e.g., employment ID, Facebook) than a police department’s usual source for photo-lineups (e.g., mugshots), and this can make the suspect’s photo stand out. This type of discrepancy will require either careful electronic editing or perhaps going to the same source as the source of the suspect’s photo (e.g., Facebook) to find fillers.

It is not the case that every aspect of background, clothing, and other features has to be exactly the same. The critical issue is whether the suspect stands out. So, for example, if the background of every photo was different from every other photo in the lineup, then perhaps the suspect’s photo would not stand out any more than any other lineup member’s photo stands out.

Articles of clothing (e.g., hats, shirts, etc.), however, could be a special concern. Specifically, if there are reasons to believe that the clothing worn by the suspect is similar to that worn by the culprit, then (a) every filler needs to be clothed that same way, or (b) the suspect needs to have his clothes changed to blend in with the fillers, or (c) all lineup member’s clothing must to be obscured from the view of the witness. Cases in which the identification of clothing is thought to have potential probative value, a separate identification procedure involving only the
articles of clothing (a clothing lineup) can be conducted independently of the identification of the suspect (see Lindsay, Wallbridge, & Drennan, 1987; Lindsay et al., 1994).

**The Single-Suspect and Minimum of Five Fillers Requirement**

A central feature of this recommendation is that the lineup should have only one suspect. There are several reasons why there should be only one suspected person in the lineup with the remainder having the status of being known-innocent fillers. Suppose, for example, a lineup were composed entirely of suspects and no fillers. All-suspect lineups have been likened to a multiple-choice test in which there is no wrong answer (Wells & Turtle, 1986; Wixted & Wells, 2017). The value of having known-innocent fillers is that unreliable eyewitnesses are likely to err on a filler rather than on an innocent suspect (assuming that the lineup is composed of good fillers).

Consider again Table 1 of the current article in which we displayed the outcomes of lineups in actual cases. These were all single suspect lineups in which the suspect was embedded among fillers. Notice that these eyewitnesses to serious crimes identified fillers approximately 37% of the times that they make an identification. If everyone in those lineups had been a suspect, all 37% of these would be mistaken identifications of innocent suspects who would then be subject to arrest and possible prosecution. These field data reinforce the dangers of having multiple suspects in a lineup that were first documented over 30 years ago through statistical proofs using data from eyewitness identification experiments (Wells & Turtle, 1986).

The recommendation that there should be at least five known-innocent fillers (thereby creating a six person lineup) for the one suspect is somewhat arbitrary. At a theoretical level, we can say that an innocent suspect is better protected from mistaken identification with a six-person lineup than a five-person lineup, which is better protection than a four-person lineup, and so on, as long as other things are equal (e.g., how good the fillers are). However, there are diminishing
returns (in terms of restricting mistaken identifications of innocent suspects) from increasing lineup size. After all, in terms of protecting an innocent suspect, an increase from two fillers to three fillers is greater than an increase from five to six, which in turn will have more impact than an increase from seven to eight fillers.

Some jurisdictions in the U.S. use more than six. Australia uses lineups that vary from eight to 10 members, and England and Wales also use more than six members. Of course, as noted in the previous discussion on selecting fillers for lineups, it is not the nominal size of the lineup that matters so much as the functional or effective size of the lineup. A lineup of 12 people would be less effective than a lineup of six people if the 11 fillers in the 12-person lineup did not fit the description of the suspect whereas the five fillers in the six-person lineup did fit the description. Of course, a 12-person lineup in which all the fillers were good fillers would provide more protection for the innocent suspect than would a six-person lineup in which all the fillers were good fillers. Although one eyewitness identification researcher has argued strongly for large increases in photo-lineup sizes to as high as 120 (Levi, 2011), at this point, we are not yet convinced that large increases in lineup size are warranted in practice. Among other things, all but one of the lineup members must be a-priori cleared as possible suspects so that they can have the definitive status of known-innocent fillers. Establishing the innocence of a large number of fillers is not an issue in a lab experiment, but it is in actual practice. As lineup size increases, it is increasingly difficult to find fillers who properly fit the description of the culprit. In addition, there are concerns about potential loss of hits (failure to identify the culprit) if a lineup becomes too large.

A Theoretical Note About Lineup Fillers
There is currently debate in the eyewitness identification literature about how good lineup fillers manage to improve overall lineup performance (i.e., how do they reduce mistaken identifications of innocent suspects more than they interfere with identifications of the culprit?). Some have suggested that the use of good fillers helps witnesses decide which facial features are diagnostic (e.g., see Colloff et al., 2016; Wixted & Mickes, 2014) whereas others have argued that good fillers simply siphon false positive identifications away from the innocent suspect more than they siphon from the culprit (e.g., see Wells, Smalarz, & Smith, 2015; Smith, et al., 2016). It is possible that both of these processes are involved and these might not be the only two possibilities. Regardless of which processes underlie the contribution that good fillers make to improved witness accuracy, the answer to this theoretical question might help clarify the best strategies for choosing fillers for lineups, and should be considered in future research.

**Recommendation 5: Pre-Lineup Instructions Recommendation**

*When inviting an eyewitness to attend a lineup procedure (photo-lineup or live lineup), police should not suggest that the suspect that will be in the lineup has been arrested or that the culprit will be present in the identification procedure. The eyewitness should be instructed that (a) the lineup administrator does not know which person is the suspect and which persons are fillers, (b) the culprit might not be in the lineup at all, so the correct answer might be not present or none of these, (c) if they feel unable to make a decision they have the option of responding don’t know, (d) after making a decision they will be asked to state how confident they are in that decision, and (e) the investigation will continue even if no identification is made.*

This recommendation addresses the concern that, if witnesses approach the identification test with the mistaken assumption that the culprit must be present in the lineup, witnesses may be predisposed towards making a positive identification. Intuition would suggest that many
witnesses are likely to presume – based on the invitation to view a lineup – that the police must have a strong suspect and, thus, their task is to determine which lineup member the suspect is. There are various strands of evidence suggesting that witnesses do make this assumption. For example, in one study, 90% of a large sample of witnesses indicated immediately after making their identification decision that they had expected the culprit to be present in the lineup and believed their task was to identify him or her (Memon, Gabbert, & Hope, 2004). In another study witnesses either viewed a lineup containing the culprit or a lineup with the culprit removed but not replaced by another filler (Wells, 1993). In the former condition, 54% of witnesses picked the culprit, 25% picked one of the fillers and 21% made no choice. When the culprit was removed but not replaced, it might be expected that around 75% of witnesses (i.e., 21% + 54%) would make no choice. Instead, only 32% of witnesses made no choice, with 68% distributing their choices across the various fillers. This finding suggests that witnesses are predisposed towards making a positive identification, though not necessarily with high confidence, provided some lineup member appears to be a reasonable match to their memory.

The most compelling evidence of the usefulness of this recommended instruction comes from studies comparing the identification performance obtained when witnesses are instructed that the culprit may or may not be in the lineup to identification performance obtained under conditions where no explicit warning that the culprit may not be present is provided. The former condition has typically been referred to as an unbiased instructions condition; the latter has generally been labeled as biased instructions. The way in which these two conditions have been enacted has varied. For example, unbiased instructions have often simply involved providing the culprit may or may not be present warning, although sometimes this has been accompanied by an instruction to the effect that there is no need to pick anyone or even an instruction emphasizing
that the consequences of a wrong decision may be dire. Biased instructions may also take many forms, such as failing to forewarn the witness that not choosing from the lineup is a perfectly appropriate response or by strongly implying or even stating that the culprit is in the lineup. For example, asking the witness to pick which lineup member is the culprit strongly implies that the witness is expected to make a positive identification decision.

Several features of the empirical findings on the effects of biased versus unbiased instructions warrant mention. First, the findings of two major reviews using meta-analytic procedures (Clark, 2005; Steblay, 1997) demonstrate that witnesses were more likely to make a positive identification decision when the lineup instructions were biased (i.e., no warning regarding possible absence of the culprit) than when they were unbiased. Second, although the magnitude of this effect varied across the reviewed studies, presumably depending on the conditions at memory encoding and the identification test, the increased likelihood of choosing was reflected in increased positive identifications from both culprit-absent (Clark, 2005; Steblay, 1997) and culprit-present lineups (Clark, 2005). Thus, it had both positive and negative effects. Third, subsequent to the publication of these reviews, findings from two studies with very large sample sizes have reinforced the impact of biased versus unbiased instructions with both adult and child witnesses (Brewer & Wells, 2006; Keast, Brewer, & Wells, 2007). Biased instructions contributed to higher rates of both mistaken identifications and correct identifications, suggesting that a failure to warn witnesses that the culprit may not be present in the lineup contributes to witnesses being prepared to accept less evidence (i.e., setting a less strict decision criterion) for making a positive identification decision. This could be a virtue if the base rate of guilty suspects were high, but it would be a vice if the base rate of guilty suspects were low, as may be often true of real police police lineups (e.g., Wixted et al., 2016).
The precise impact of witnesses lowering their decision criterion will also depend on factors, such as the characteristics of the various lineup members (Brewer, Weber, & Semmler, 2005). For example, if the lineup is biased against the suspect by virtue of the suspect being the only plausible lineup member, a lower decision threshold would increase the likelihood of (a) a correct identification if the suspect is the culprit and (b) a mistaken identification of an innocent suspect if the suspect is innocent. In contrast, if the suspect is presented in an array of highly plausible fillers, instructions that lead witnesses to set a lower decision threshold may lead to responses being spread more evenly across all lineup members.

It is important to note one qualification of the pattern of findings typically found when contrasting the impact of unbiased versus biased instructions. The effect of unbiased instructions may be negated if the witness receives an explicit suggestion prior to viewing the lineup that the culprit may be present in the lineup. In a study in which witnesses were presented with culprit-absent lineups only, witnesses received the instruction surely you are going to be able to pick the person out from the lineup prior to receiving the instruction that the culprit may or may not be present in the lineup (Quinlivan et al., 2012). Witnesses who received that suggestion, followed by unbiased instructions, were three times more likely to identify the suspect (who, in this study, was innocent) than were witnesses who received unbiased instructions without any prior suggestive comment. Witnesses in the former condition were also more confident in the accuracy of their erroneous identification decision than those in the latter condition.

The research findings on instructing witnesses prior to their viewing a lineup have clear implications. First, when inviting an eyewitness to attend a lineup procedure, police should not suggest that a suspect has been arrested or that the culprit will be present in the identification procedure. Second, in our experience some witnesses seem to be under the misconception that
the investigation hinges on their identification decision. Consequently, witnesses should also be
told that the investigation will continue even if no identification is made.

Third, it should be made quite clear to the witness that the culprit may or may not be in the
lineup and that they do not have to select any of the lineup members. In other words, responses
such as not present or none of these are quite appropriate. This is because, in our experience,
witnesses are often under the impression that they have to choose someone from the lineup.

Fourth, to ensure that the witness does not lose sight of the fact that such response options are
appropriate, there should be an explicit not present response option accompanying the lineup
members from which eyewitnesses can choose. In the case of a photo-lineup this option may be
located below the array of lineup faces. In the case of a live lineup a response sheet that shows
the possible response options can be used: lineup member numbers (i.e., 1, 2, … 6), not present,
and don’t know.

Finally, lineup administration procedures should accommodate the possibility that the
witness may look at the lineup and be unwilling to pick someone or to respond not present
because, for example, they cannot decide between two or more lineup members or they are
uncertain about whether the culprit is in the lineup. For that witness (i.e., one who really has no
idea about what to do), an appropriate response may be to say don’t know rather than not
present. Both adult and child witnesses use options such as don’t know or not sure when they are
made explicit, with the frequency with which they are used varying considerably depending on
the encoding stimuli and lineup materials (Brewer, Keast, & Sauer, 2010; Perfect & Weber,
2012). Moreover, there is some evidence indicating that positive identifications of a suspect that
were made when there was an available don’t know option were more diagnostic of suspect guilt
than when no such option existed (Perfect & Weber, 2012).
The availability of an option to respond don’t know is likely to reduce the likelihood of low confidence positive identifications, which research shows are often inaccurate (Brewer & Wells, 2006; Wixted & Wells, 2017). Although there are strong grounds for always questioning the reliability of low confidence identifications, there may be a tendency on the part of police or prosecutors to argue, for example, that the initial low confidence identification was made from a photo that was not a good likeness to the suspect, thereby purportedly providing a reasonable explanation for the witness’s low confidence. Further, although mock-jury studies have shown that, when cross-examination highlights how a witness expressed much lower confidence in their identification of the defendant at the time of the initial identification than they did in court, mock-jurors (appropriately) downgrade the credibility of the witness and the culpability of the defendant (Bradfield & McQuiston, 2004; Jones, Williams, & Brewer, 2008), this pattern does not always occur (e.g., Douglass & Jones, 2013). For example, Jones et al., 2008, if a witness expressed high confidence in their identification decision during a trial despite a low confidence initial identification, jurors tended to excuse the confidence inflation if the witness appeared to have experienced an epiphany about their initial low confidence identification (e.g., “I wasn’t very confident at the time of the identification because I was scared back then.”). As long as these kinds of excuses for initial low-confidence identifications are permitted, we argue that it is crucial that there should be an explicit don’t know response option, which can be located alongside the not present option.

As a final point on instructions, we note that some jurisdictions have used what has been referred to as an “appearance change” instruction. This instruction was among a set of guidelines developed by a US Department of Justice working group on the collection of eyewitness evidence; it recommended that, prior to being shown a lineup, eyewitnesses should be told that
“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). We have not included this instruction in the current set of recommendations because subsequent research has shown that the instruction increased false identifications but did not increase culprit identifications (Charman & Wells, 2007; Molinaro, Arndorfer, & Charman, 2013).

**Recommendation 6: Obtain Immediate Confidence Statement Recommendation**

*An confidence statement be taken from witnesses as soon as an identification decision is made.* Note that a confidence statement should be recorded if a witness positively identifies someone or if a witness indicates the culprit is not present. As an example, the confidence statement could be collected as a numeric response (i.e., on a scale from 0% confident to 100% confident). Alternatively, confidence could be collected using a verbal scale (e.g., “positive,” “probably,” “maybe”). If neither scale is used, we recommend a verbatim record of whatever verbal statement exists, not a summary or paraphrase. If the witness’s response is “don’t know,” a confidence statement should be recorded if the witness spontaneously provides one. Otherwise, no confidence statement should be solicited for “don’t know” responses. The pre-lineup instructions should communicate that a confidence statement will be recorded to prevent witnesses from drawing erroneous conclusions that their confidence is only being assessed because the lineup administrator thinks their decision is incorrect (see Recommendation 5).

This recommendation is based on the fact that eyewitness confidence is a useful cue to the accuracy of a witness’s decision, particularly when instructions do not imply the presence of a culprit (Quinlivan et al., 2012), double-blind administration is used, and fair lineups are presented (Wixted & Wells, 2017). Specifically, confidence predicts accuracy among witnesses
who choose from a photospread (e.g., Brewer & Wells, 2006), among those who provide immediate confidence reports (cf., those whose reports were withheld for as little as 5 minutes, Brewer, Keast & Rishworth, 2002), and when witnessing conditions vary widely (e.g., Lindsay, Read, & Sharma, 1998).

The corpus of data suggesting that confidence is a useful cue to eyewitness accuracy stands in stark contrast to DNA exonerations in which innocent people were mistakenly identified by highly confident eyewitnesses, most of whom were demonstrably less confident at the time of the initial identification (e.g., Garrett, 2011). If confidence is a useful cue to accuracy, how could these mistaken witnesses have been so confident? As described below, the answer lies in the fact that intervening post-decision events can dramatically shift witnesses’ reports of their confidence from low at the time of identification to high at the time of trial. As a result, these shifts render delayed reports nearly useless as cues to accuracy. However, if confidence reports are taken immediately after an identification decision, the integrity of confidence as a cue to accuracy is enhanced considerably.

We focus on immediate confidence reports because seemingly innocuous post-decision events can profoundly distort witnesses’ confidence reports, undermining what could have been forensically meaningful information from an eyewitness. One of the most heavily researched of these events is post-identification feedback in which a simple comment confirming the witness’s decision dramatically inflates confidence reports. In the original test of the post-identification feedback effect (Wells & Bradfield, 1998), researchers told inaccurate witnesses: “Good, you identified the suspect.” That simple comment resulted in 50% of inaccurate witnesses reporting that their confidence was a 6 or 7 on a 7-point scale (compared with 15% of witnesses in the control condition). Importantly, the inflated confidence report created by feedback is a
**Retrospective Judgment** because witnesses indicate how confident they were at the time of their identification, before they knew their decision was correct. Any resulting confidence inflation obviously obscures a true picture of the witness’s experience at the time of the identification decision. The effect of post-identification feedback is robust and reliable (see meta-analysis by Steblay, Wells, & Douglass, 2014). It also features prominently in some judicial decisions as courts grapple with how to ensure that eyewitness identification testimony truly reflects the witness’s experience of making an identification decision, rather than the influence of extra-memorial variables (e.g., *New Jersey v. Henderson*, 2011; *Oregon v. Lawson*, 2012).

Beyond affecting witnesses’ self-reports, distorted confidence judgments complicate assessments of witness identification decisions. Indeed, several experiments show us that evaluators rate inaccurate witnesses who have received confirming feedback as more credible than those who received disconfirming feedback or no feedback, even when evaluators are instructed to ignore the feedback and even when an explicit confidence statement is not available (Douglass, Neuschatz, Imrich, & Wilkinson, 2010). Because post-identification feedback inflates the confidence of inaccurate witnesses more than the confidence of accurate witnesses, it also impairs evaluators’ ability to distinguish accurate from inaccurate witnesses (Bradfield, Wells, & Olson, 2002; Smalarz & Wells, 2014b).

Recommending an immediate confidence report is an important companion to the recommendation that the entire lineup procedure be video-recorded (see Recommendation 7). Not only can video-recording demonstrate whether post-identification feedback was given, but it can also preserve witness non-verbal cues that may signal accuracy (Matuku, Douglass, & Charman, 2018). In the absence of a video-recorded identification procedure, triers of fact who learn that a witness’s confidence has inflated over time are generally unwilling to impugn the
witness’s credibility (Bradfield & McQuiston, 2004), especially if a compelling explanation accompanies the inflation (e.g., ‘I had an epiphany!’ Jones et al., 2008). However, if evaluators see the identification procedure in which a witness’s initial confidence is lackluster, their assessments of a highly confident trial witness are less positive, an important shift when the highly confident witness has identified the wrong person (Douglass & Jones, 2013).

Beyond distorting evaluators’ ability to assess witnesses, contaminated witness reports may also unduly shape preliminary investigations by (a) triggering biased evaluations of subsequent pieces of evidence and/or (b) biasing the integration of evidence against the identified suspect. For example, a witness who is highly confident in a mistaken identification may trigger investigators to view the suspect’s alibi as weaker than it would have been otherwise or may suggest to investigators that they suspend pursuit of additional suspects (for a discussion of these effects, see Charman, Douglass, & Mook, in press). Recording a witness’s immediate confidence eliminates the potential for subsequent inflations to go unnoticed.

If lineup administrators follow the recommendation to conduct double-blind procedures (Recommendation 3), they will be unable to provide post-identification feedback because they will not know which person is the suspect. Therefore, they will be unable to confirm (or disconfirm) any decision made by the witness. However, even vague positive comments can be interpreted by witnesses as confirming feedback (Dysart, Lawson, & Rainey, 2012). Therefore, consistent with the recommendation on pre-lineup instructions (Recommendation 5), witnesses should be explicitly told that the lineup administrator does not know which person is the suspect and which people are fillers.

Even if administrators correctly implement double-blind procedures and pre-lineup instructions, it is still imperative to collect immediate confidence reports because witnesses may
encounter post-identification feedback in multiple other ways. For example, the mere fact that a case proceeds to trial is a form of confirmation that a prior identification was correct. Another potential source of feedback comes from witnesses who conduct their own investigations via social media searches, often with co-witnesses who may be able to “confirm” that an Instagram or Facebook photo represents the culprit (Douglass & Smalarz, in press).

We recommend that confidence statements be collected on a graded scale using words (e.g., “positive,” “probably,” “maybe”) or numbers (e.g., from 0% confident to 100% confident). The key element of this recommendation is that an immediate record of a witness’s confidence is collected. Immediate confidence estimates are the only way to ensure that post-identification variables do not contaminate subsequent confidence reports.

**Recommendation 7: Video-recording Recommendation**

*The entire identification procedure, including pre-lineup instructions and witness confidence statement, should be video-recorded.*

To preserve a faithful record of the conditions under which witnesses make their identifications, we recommend that the entire interaction between the police and the witness be video-recorded from time when the witness is given the very first pre-lineup instructions through to the completion of the procedure when witnesses have provided their confidence statements. The video must capture all the features of the administration, including the interactions among the lineup administrator, the witness, and the lineup members (whether they are presented live or by photo). Under some circumstances, fully capturing the lineup administration may necessitate cameras recording from multiple angles. Although the authors of the original scientific review paper (Wells et al., 1998) recognized some of the benefits of video-recording a lineup administration, they did not include video-recording as one of their recommendations at that
time. In part, video-recording of the identification procedure was left out of the original set of recommendations because of concerns about the costs associated with making the record, including costs for equipment and materials. Since that time, the cost of video-recording interactions has decreased considerably and most adults have cellular phones capable of rendering high-quality video-records. Thus, we believe it is time to make the video-recording of lineup administrations standard practice as it is in some jurisdictions in other countries (e.g., South Australia).

There are a variety of reasons why it should be standard practice to implement video-recording of identification procedures. Kassin (1998) noted two reasons why video-recording identification procedures should be considered best practice. First, police reports of what happened during an identification procedure may be incomplete or even inaccurate given that they are based on officers’ recollections of what happened during the procedure - recollections that are subject to the typical foibles of human memory. Although there is no direct evidence about the accuracy of police reports of identifications, police reports of interrogation procedures omit important details about the procedures used (Kassin, Kukucka, Lawson, & DeCarlo, 2017). Moreover, testimony from witnesses about what happened during the procedure is even more likely to be subject to error given that police reports are more likely to be written shortly after the event, whereas testimony may be given months or years later. Video-recording the procedure eliminates the need to rely on police officers’ or witnesses’ memories for the event. Second, video-recording the procedure allows for a complete record of the event to be available for judges, attorneys, and jurors to view at a later time, allowing them to evaluate the procedure for suggestiveness and to see for themselves how the witness reacted to the task (including the time it took them to make an identification, if any, and any non-verbal behaviors associated with their
expressions of confidence). It is also possible that video-recording of the identification procedure could encourage administrators to carefully adhere to best practices and deter them from engaging in any suggestive practices (Kassin, 1998). In the remainder of this section, we expand upon each of these justifications for video-recording identification procedures, citing relevant research when it is available.

First, video-recording identification procedures has the benefit of providing a more precise and accurate accounting of what happened during the procedure. Although it might be a rare occurrence, video-recording the procedure will make it more difficult for police officers to intentionally fabricate their reports of what occurred during the lineup administration. There is evidence of some police officers purposefully misrepresenting case-related events (Orfield, 1987; 1992; Slobogin, 1996), but even in the case of a conscientious, well-meaning officer, there are benefits of a more accurate reporting of the procedure through video-recording. DFor example, the memories of police officers are subject to the same cognitive errors as are those of others, including interference (e.g., Kane & Engle, 2000; Nairne, 2002) and memory intrusions from scripts (e.g., Greenberg, Westcott, & Bailey, 1998; Kleider, Pezdek, Goldinger, & Kirk, 2008). One can easily imagine that memories from other lineups conducted or scripts for what should have happened might interfere with what police officers remember—and consequently report—about an identification procedure. Indeed, there have been cases in which the police officer reports reading mandated instructions verbatim to the witness, when the video-record reveals ad-libbing on the part of the officer, improvisation that introduced suggestiveness into the procedure. Whether or not the officer was attempting to intentionally mislead or was genuinely mistaken, the actual procedure could be easily reviewed if it is video-recorded, and fallibility of memory is no longer an issue.
In contrast, witnesses’ memories for a procedure are not likely to be influenced by interference or scripts given most witnesses have never participated in a lineup before. However, their lack of expertise with the procedure will likely result in reports that are less complete (e.g., Chase & Simon, 1973; Gobet & Simon; 1996). Moreover, their reports of the procedure may be affected by the decay of their memory trace over time (Deffenbacher, Bornstein, McGorty, & Penrod, 2008) or post-event information (Loftus, 2004). The video-record would also provide clear information about how long it took the witness to make an identification. Witnesses’ estimates of time are often inaccurate (Yarmey, 2000), but time-to-identification is an important postdictor of witness accuracy, with shorter times taken to make the identification more likely to result in an accurate identification (Sporer, 1992). Although research has not provided a definitive cutoff that allows us to discriminate between accurate and inaccurate identifications (Weber, Brewer, Wells, Semmler, & Keast, 2004; Brewer, Caon, Todd, & Weber, 2006), having an accurate record of the witnesses’ times-to-identification can provide useful information to fact finders. In sum, despite knowing no evidence directly examining the accuracy of police officers’ and witnesses’ memories of identification procedures, we are confident that the basic cognitive research on memory errors generalizes to this context. Thus, we recommend video-recording of the lineup procedure to provide a faithful record of what happened during the lineup administration.

The second benefit of video-recording the procedure is that judges, attorneys, and jurors can view the evidence of the lineup procedure when they are faced with making legal decisions related to the identification. When deciding a motion to suppress the identification, the judge could review the video and evaluate the suggestiveness of the procedure herself rather than relying on attorneys’ characterizations of the procedure based on their readings of police reports.
and witness testimony. Attorneys might review the video-record when deciding whether to offer a plea or encourage a client to accept one. Moreover, as a suspect does not have the right to have an attorney present at a pre-indictment identification procedure (United States v. Ash, 1973), the video-record would allow defense attorneys to view any pre-indictment procedures when preparing to defend their clients who had been identified. Finally, the record could be introduced as evidence at trial so that fact-finders can judge for themselves whether the procedure was suggestive or whether the witnesses engaged in any behaviors that either enhanced or diminished their credibility.

Although the data on this topic are limited, and there are no known data on how judges or attorneys might evaluate video-records of identification procedures, there are a handful of studies that explore the effects of watching the video-record of an eyewitness identification procedure on mock juror judgments. Most of these studies have tested whether watching the video-record of the identification procedure helps jurors judge the accuracy of witness identifications. In one study, participants who viewed both witness testimony as well as the video-record of the identification procedure were better able to distinguish between accurate and inaccurate witnesses than were participants who saw only the witness testimony (Reardon & Fisher, 2011). In another study, participants judged the accuracy of both accurate and inaccurate identifications after watching only the witness testify about making an identification, only the identification procedure, or both the testimony and the identification procedure (Beaudry et al., 2015). One of three procedures had been used to obtain the identification: (a) the administrator did not know which lineup member was the suspect (double-blind procedure), (b) the administrator knew which lineup member was the suspect (single-blind), or (c) a single-blind procedure plus confirmatory feedback when the witness identified the suspect. Confirmatory
feedback interfered with the ability of participant-judges to differentiate between accurate and inaccurate witnesses, with participants judging the witnesses who received confirmatory feedback to be more accurate irrespective of their actual accuracy, probably because these witnesses also reported higher confidence in the accuracy of their identification. Participants’ judgments of accuracy were unaffected by whether the lineup was conducted using single- or double-blind procedures, irrespective of whether the participants had viewed the video of the procedure. However, these videos were relatively short ($M = 1m, 37s$) and it is not clear from the report of the study how much influence the administrators exerted in the single-blind conditions. Thus, the limited evidence is mixed as to whether a video-record helps lay people better judge the accuracy of an identification.

There is also mixed evidence regarding whether a video-record helps laypersons better judge the accuracy of an identification. In one series of studies, the verdicts of mock jurors were unaffected by hearing evidence that a lineup administrator knew which lineup member was the suspect (Wright, Carlucci, Evans, & Compo, 2010). In another study, however, watching a video-recording of a single-blind versus a single-blind procedure did influence jurors’ verdicts (Modjadidi & Kovera, 2018). In that study, mock jurors recruited from the community watched a trial that contained testimony that the police officer who conducted the lineup procedure had or had not known which lineup member was the suspect (single- vs. double-blind administration). Half the participants also saw a video-recording of a simulated identification procedure, with no suggestive administrator behavior in the double-blind version and administrator behaviors suggesting the suspect in the single-blind version). Although participants were sensitive to the potential suggestiveness of the single-blind procedure without the video-record (unlike in the Wright et al. studies), watching the video-recording increased their ratings
of procedural suggestiveness and decreased their guilty verdicts when the video-recording depicted the single-blind administration. Thus, the limited evidence available suggests that watching the video-recording of a lineup administration will help jurors (and perhaps judges and attorneys) evaluate whether a lineup procedure was suggestive.

Finally, video-recording the identification procedure could encourage administrators to ensure that their conduct conforms with best practices and deter them from engaging in any suggestive procedures. Although we know of no studies that directly test whether video-recording will change the behaviors of administrators of identification procedures, the issue has been examined in the context of interrogation practices. In a mock crime and interrogation paradigm, half of the participant-suspects committed a mock crime and half did not; all were interrogated by police officers (Kassin, Kukucka, Lawson, & DeCarlo, 2014). Half of those police officers were told that the interrogations were being recorded, half were not even though all the interrogations were recorded. Police officers who knew that they were being recorded were significantly less likely to use minimization tactics and somewhat less likely (although not significantly so) to use maximization tactics; both tactics are known to increase rates of false confessions. In addition, the participant-suspects, who did know about the video-recording manipulation, reported that the police officers who did not know they were being recorded tried harder to make them confess to the crime than did police officers who knew about the recording. Thus, there is evidence, albeit from outside the eyewitness arena, that video-recording police procedure can deter undesirable behaviors.

In sum, both logic and the available empirical evidence support the recommendation to video-record all identification procedures. If the procedure is to be recorded, it should be recorded in a way that captures all relevant information about the procedure, including the verbal
and non-verbal behaviors of the witness, the administrator, and the lineup members. Research from the interrogations literature makes it clear that camera angle matters, in that people attribute causality to the person who is the focus of the video-record (Lassiter, 2010; Lassiter, Ware, Ratcliff, & Irvin, 2009; Ratcliff, Lassiter, Schmidt, & Snyder, 2006). In the case of interrogations, a camera focused on the suspect causes evaluators to be more likely to view a confession as voluntary and more likely to judge them to be guilty than if the camera was focused on the interrogator or equally on the suspect and the interrogator. It is reasonable to believe that cameras focused on witnesses may cause evaluators to overlook suggestive behaviors on the part of administrator or features and behaviors of suspects that make them stand out from the other lineup members. However, a video-record of an identification procedure that includes information about the witness, administrator, and lineup members can provide a complete record of the procedure that documents suggestive practices when they are present and protects the police from unjustified and time consuming claims of bias when the procedure was free from suggestion.

Recommendation 8: Avoid Repeated Identification Procedures

Using the Same Witness and Suspect Recommendation

Repeating an identification procedure with the same suspect and same eyewitness should be avoided regardless of whether the eyewitness identified the suspect in the initial identification procedure. This recommendation holds no matter how compelling the argument in favor of a second identification might seem (e.g., the original photo of the suspect was not as good as it could have been; the witness was nervous on the first identification test and is calmer now; the witness had an epiphany after the first test and now clearly remembers the perpetrator; the
witness admits to not having been honest on the first test and is ready to be honest now, etc.).
The importance of focusing on the first identification test cannot be emphasized strongly enough.

There are certain kinds of forensic evidence for which repeated testing is not only acceptable but also desirable. A crime scene fingerprint comparison to a suspect’s prints, for example, might be subjected to repeated comparisons to confirm a match conclusion. Likewise, given a sufficient amount of blood or semen from a crime scene, forensic examiners might test only a small portion of the sample for DNA, allowing a second or third forensic analyst to repeat the test procedure independently on the remainder of the sample. However, eyewitness identification evidence has a unique characteristic that makes it unsuitable for what might be called “repeated testing.” Whether the eyewitness is asked to make an identification with a showup or a lineup, there is only one uncontaminated chance for any given eyewitness to make an identification of a particular suspect. Any subsequent identification test with that same eyewitness and that same suspect is contaminated by the eyewitness’s experience on the initial test.

For purposes of our recommendation, repeated identification tests refer to a situation in which an eyewitness is given a second identification test (or more) with the same suspect that appeared in an earlier identification test. We are not referring to a situation in which an eyewitness is given an identification test with one suspect and then, after rejecting that suspect is given an identification test with a different suspect. Nor are we referring to a situation in which there were multiple perpetrators and the eyewitness is given a separate identification test for each suspect.

There are many ways in which the use of repeated ID procedures surface in criminal cases. For example, eyewitnesses might view a mugbook that contains the suspect prior to
viewing a lineup that includes that suspect. Another example of a repeated identification is when the eyewitness first views a showup and then later confronts a lineup with that same suspect. It is also a repeated identification when the eyewitness views a lineup and makes a tentative pick (or no pick) and then, at a later time, sees a second lineup with that same suspect and new fillers. Sometimes a witness is first shown a photo-lineup, and then is later shown a live lineup using that same suspect. Other times the first lineup uses one photo of the suspect, and the second lineup contains a different photo of that same suspect. Perhaps the most common repeated identification procedure of all is when the witness made an earlier out-of-court identification (from a showup or a lineup), and then is asked to repeat that identification in court at pre-trial hearings or at trial (i.e., an in-court or “dock” identification).

At a theoretical level, there are at least three processes by which an initial identification test that includes a given suspect can contaminate a second identification test if the second test includes that same suspect (Deffenbacher, Bornstein, & Penrod, 2006; Steblay & Dysart, 2016). One such process is memory-source error (or “source monitoring error”, see Johnson, Hashtroudi, & Lindsay, 1993). Memory-source error is implicated when the eyewitness perceives the suspect in the second identification procedure to be familiar and misattributes the familiarity to the original witnessed event rather than to the fact that the previous identification procedure included that face. In this type of effect, even though the eyewitness did not identify the suspect in the first identification procedure, that person’s face was made familiar by its appearance in the first procedure. This source misattribution effect, which involves a dissociation between familiarity and an awareness of the source of that familiarity, was first demonstrated over 40 years ago (Brown, Deffenbacher, & Sturgill, 1977). It is closely related to other phenomena such as familiarity without awareness (Mandler, 1980), the false fame effect (Jacoby,
Kelley, Brown, & Jashecko, 1989), and the “bystander effect,” in which a bystander to the crime is identified as the perpetrator due to a misattribution of familiarity (Ross, Ceci, Dunning, & Toglia, 1994). Simply being exposed to an innocent suspect in a mugbook, showup, or a lineup increases the chances of that person being identified in a later lineup even if the witness did not choose the person in the first identification procedure (e.g., Brown et al., 1977; Haw, Dickinson, & Meissner, 2007; Hinz & Pezdek, 2001; Lawson & Dysart, 2012; Steblay, Tix, & Benson, 2013).

A second process by which the first identification procedure can contaminate the second identification procedure is when the eyewitness identifies the suspect in the initial identification procedure and is later given a second identification procedure with that same suspect and a different set of fillers. In this case, the initial identification, even if mistaken, causes the witness to simply repeat the same identification in the second identification procedure. A meta-analysis has provided strong evidence that a mistaken identification in an initial identification procedure tends to be repeated in a second identification procedure that contains the mistakenly identified person (Deffenbacher et al., 2006); more recent studies have provided additional support for this effect (e.g., Godfrey & Clark, 2010; Goodsell, Neuschatz, & Gronlund, 2009; Hinz & Pezdek, 2001; Lawson & Dysart, 2012; Pezdek & Blandon-Gitlin, 2005; Steblay, Tix, & Benson, 2013; Valentine, Davis, Memon, & Roberts, 2012). Researchers have argued that commitment, which is the powerful tendency to stick with an earlier decision that was freely made, a process closely related to cognitive dissonance (e.g., Knox & Inkster, 1968), is the psychological mechanism underlying this effect. But, commitment is not the only process by which an initial mistaken identification results in a repeat mistaken identification of the same person in a subsequent identification test. For example, there is evidence that the act of identifying an innocent person in
an initial identification procedure changes the eyewitness’s memory away from the culprit and toward the person identified, a process that is intensified if the witness is given confirming feedback following the initial mistaken identification (Smalarz & Wells, 2014). Moreover, in the basic literature on testing effects, there is good evidence that alternatives that are incorrectly selected on an initial multiple-choice test are incorrectly recalled as being the correct answer on subsequent tests (Roediger & Marsh, 2005; Butler, Marsh, Goode, & Roediger, 2006; Marsh, Roediger, Bjork, & Bjork, 2007).

Although not yet specifically tested in controlled experiments, there is a third possible process by which repeating the same suspect in a second identification procedure can contaminate the second identification decision. Specifically, it is possible for the eyewitness who makes no identification in the first procedure (e.g., a photo-lineup) to consciously and explicitly recognize that there is only one person in the second procedure (e.g., a live lineup) who was also in the first procedure. In effect, this “gives away” the hypothesis of the police investigators by communicating to the eyewitness which person is the suspect (i.e., the person in common between the two procedures). In other words, this type of repeated procedure manages to violate a fundamental characteristic of a good lineup, namely that there not be aspects of the procedure that leak information about which person is the suspect in the lineup (Wells & Luus, 1990).

Any of these processes might occur in a given situation that involves repeating the suspect in an identification procedure. For this reason, we recommend that repeated identification procedures be avoided. Of course, it could be argued that if the witness identifies the suspect in the first identification procedure then there is no harm in repeating the identification procedure. After all, the identification has already been made and the second identification procedure is merely a confirmatory process. However, there is good evidence that
repeated testing of eyewitnesses leads to artificially elevated levels of eyewitness confidence (Shaw & McClure, 1996). Repeatedly asking a question appears to lead to increased ease or fluency of answering the question, which leads witnesses to develop a stronger sense of confidence that their answer is correct, even when their answer is incorrect (Shaw, 1996).

**Recommendation 9. Showups Recommendation**

Showups should be avoided whenever it is possible to conduct a lineup. Cases in which it is necessary to conduct a showup should use a number of the procedural safeguards that are used for lineups, including the elimination of suggestive cues, a warning that the detained person might not be the culprit, video-recording the procedure, and securing a confidence statement.

The term *showup* refers to an eyewitness identification procedure in which a single individual is presented to the eyewitness and the eyewitness is asked whether this is the person who committed the crime in question. The primary defining feature that separates a showup from a lineup is the use of fillers; a showup has no fillers, whereas a lineup does. Showups have long been heavily criticized as being extremely suggestive, a criticism that goes back more than 100 years (Gross, 1911). The U.S. Supreme Court, in reference to a showup identification, said in 1967 that “It is hard to imagine a situation more clearly conveying the suggestion to the witness that the one presented is believed to be guilty by the police” (*United States v. Wade*, 1967, p. 226). Despite the strong language seemingly condemning showups, the U.S. Supreme Court has consistently supported the admissibility of showups (e.g., see *Neil v. Biggers*, 1972; *Manson v. Brathwaite*, 1977) as long as the witness shows evidence of reliability in other ways. The Court listed five factors to consider that might indicate that a suggestive procedure could nevertheless be reliable, namely the witness had a good opportunity to view the culprit, the witness paid good attention while witnessing, the witness gave a good description of the perpetrator, there was a
short delay between the witnessed event and the identification, and the witness made the identification with high confidence. These criteria have been strongly criticized by eyewitness scientists on a variety of counts, especially when the procedure was suggestive (Wells & Quinlivan, 2009).

In practice, showups tend to be conducted under specific conditions. That is, when an individual is detained who fits the description of the culprit of a crime that has just occurred in the general vicinity. As discussed in the introduction to this article, there is a rationale for sometimes permitting showups under these conditions (fit description, proximity to crime, and soon after the crime). The rationale is that there is not enough time to construct and conduct a lineup procedure because the detained person can only be detained for a relatively short amount of time unless there are grounds for arrest. Fitting a description of a culprit is not, in and of itself, grounds for arrest. So, the choice is not between conducting a lineup and conducting a showup. The choice is between conducting a showup or setting the detained person free, thereby potentially creating a public safety issue. Moreover, because showups often result in eyewitnesses rejecting the detained person as being the culprit, showups can result in innocent people being quickly excluded as possible suspects.

Experiments comparing lineups to showups reveal that lineups are clearly superior to showups in terms of the lineup procedure’s ability to distinguish between innocent and guilty suspects (e.g., see Clark, 2012; Dekle, Beal, Elliot, & Huneycutt, 1996; Gronlund et al., 2012; Mickes, 2015; Steblay, Dysart, Fulero, & Lindsay, 2003; Wetmore et al., 2015; Yarmey, Yarmey, & Yarmey, 1996). Interestingly, however, witnesses are no more likely to make an affirmative identification with a showup than with a lineup (Gonzalez et al., 1993; Wells, 2001). In fact, witnesses make more affirmative identification responses to lineups than to showups,
probably because there are more people from whom to choose. However, inaccurate identifications from showups always fall on the innocent suspect whereas in lineups the inaccurate choices can land on innocent suspects or be spread across fillers (Smith et al., 2016; Wells, 2001). In several ways, the fillers used in a lineup are safeguarding features of an innocent suspect’s protection from mistaken identification, a form of protection that does not exist for showups.

Our recommendation, however, is that showups should be avoided (and a lineup conducted instead) whenever possible. In fact, it is not uncommon for there to be probable cause for arresting a detained person instead of doing a showup. If, for example, the detained person (1) is in possession of stolen goods, (2) is in possession of an unauthorized firearm or other weapon, or (3) has an outstanding warrant, then the person should be arrested. A lineup for the arrested individual can then be arranged rather than a showup. Another situation in which showups can be avoided occurs when there are multiple witnesses. In cases of multiple witnesses, a showup can be conducted with one of the witnesses and, if an identification is made, the identification becomes grounds for arrest. The remaining witnesses can then be preserved for the more reliable lineup procedure.

Reducing the Suggestiveness of Showups

Although showups are clearly less reliable than lineups, there are some ways to reduce the suggestiveness of showups. First, many of the features of a good lineup procedure can be incorporated into showup procedures. For example, recent research shows that a pre-showup instruction warning the witness that the person they are going to be shown might not be the culprit results in a lower false identification rate with little effect on reducing accurate identifications (Eisen & Smith, 2017). An additional instruction to the eyewitness prior to the
showup, called the *additional-opportunities instruction*, can be effective in reducing mistaken identifications with little or no reduction in accurate identifications (Eisen & Smith, 2017; Smith, Wells, Lindsay, & Myerson, 2018). The additional-opportunities instruction simply tells witnesses prior to viewing the detained person that if they do not think the person is the culprit, they might have additional opportunities to view someone else. The theory behind the additional-opportunities instruction is that witnesses set a low criterion for making an identification with showups because they assume that this is their only opportunity to identify the culprit.

As with lineups, eyewitnesses who participate in a showup procedure should be asked to report how confident they are in their identification or rejection decision. In addition, as with lineups, showups should be video-recorded, a recommendation that is relatively easy to implement now that dashboard-mounted and body-worn cameras are increasingly commonplace (see Recommendation 7). The video of the showup should start before the witness is instructed and continue through the witness’s confidence statement.

Because showups are conducted in the field during which search-and detain-operations are actively unfolding, care should be taken to ensure that witnesses do not overhear police radio conversations that could prejudice a showup procedure. In an experiment using a high-reality paradigm for studying showup identifications, overhearing the suggestion that the sheriff had caught the guy significantly increased false identifications from showups but did not affect accurate suspect identifications (Eisen, Skerrit-Perta, Jones, Owen, & Cedre, 2017).

Clothing is often a central feature of an eyewitness’s description that police use for finding a person who is then detained for a showup. The concern is that the eyewitness might identify the person based largely on the clothing (see Lindsay, Wallbridge, & Drennan, 1987). In
cases where the person was detained based on a clothing description, consideration should be given to covering the person’s clothing with a blanket for the showup identification test.

For legal reasons, eyewitnesses are usually brought to the location of the detained person for a showup rather than the detained person being brought to the eyewitness because transporting the detained person is usually considered an arrest. Consistent with our recommendations, if the detained person can be arrested, then a lineup should be conducted rather than a showup. Taking the eyewitness to the detained individual rather than the reverse also makes it easier in multiple-witness cases to ensure that the witnesses are not contaminated by observing an identification decision made by another witness. As noted before, if one witness makes an identification, then that would normally be considered probable cause for arrest and the remaining witnesses can later view a lineup instead of a showup. Placing the detained person in handcuffs or in the back seat of a patrol car for the showup can suggest to the witness that the person has been arrested. This, in turn, can suggest to the eyewitness that there is evidence against the detained person beyond simply matching the description of the suspect. Hence, unless there is reason to believe that the detained person is a flight risk, these types of restraints should be avoided during the showup.
References


https://doi.org/10.1037/law0000115


https://dx.doi.org/10.3758/BF03194051


https://doi.org/10.1146/annurev.psych.53.100901.135131


https://dx.doi.org/10.2307/1599834


https://dx.doi.org/10.1002/acp.1796

https://doi.org/10.1002/acp.939


Table 1. Summary statistics on 11 published field studies of eyewitness identification.

<table>
<thead>
<tr>
<th>Authors</th>
<th>No. of possible IDs</th>
<th>IDs of suspects</th>
<th>IDs of fillers</th>
<th>Rejections (No ID)</th>
<th>Suspect%</th>
<th>Filler%</th>
<th>No ID%</th>
<th>% Making an ID</th>
<th>Suspect ID rate among all IDs</th>
<th>Filler ID rate among all IDs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behrman &amp; Davey (2001)</td>
<td>58</td>
<td>29</td>
<td>14</td>
<td>15</td>
<td>50.0%</td>
<td>24.1%</td>
<td>25.9%</td>
<td>74.1%</td>
<td>67.4%</td>
<td>32.6%</td>
</tr>
<tr>
<td>Behrman &amp; Richards (2005)</td>
<td>461</td>
<td>238</td>
<td>68</td>
<td>155</td>
<td>51.6%</td>
<td>14.8%</td>
<td>33.6%</td>
<td>66.4%</td>
<td>77.8%</td>
<td>22.2%</td>
</tr>
<tr>
<td>Horry et al (2014)</td>
<td>833</td>
<td>382</td>
<td>149</td>
<td>302</td>
<td>45.9%</td>
<td>17.9%</td>
<td>36.3%</td>
<td>63.7%</td>
<td>71.9%</td>
<td>28.1%</td>
</tr>
<tr>
<td>Horry, Memon, Wright, &amp; Milne</td>
<td>1039</td>
<td>406</td>
<td>273</td>
<td>360</td>
<td>39.1%</td>
<td>26.3%</td>
<td>34.6%</td>
<td>65.4%</td>
<td>59.8%</td>
<td>40.2%</td>
</tr>
<tr>
<td>Klobuchar, Steblay, &amp; Caligiuri (2006)</td>
<td>178</td>
<td>63</td>
<td>20</td>
<td>95</td>
<td>35.4%</td>
<td>11.2%</td>
<td>53.4%</td>
<td>46.6%</td>
<td>75.9%</td>
<td>24.1%</td>
</tr>
<tr>
<td>Memon, et al. (2011)</td>
<td>1044</td>
<td>456</td>
<td>437</td>
<td>151</td>
<td>43.7%</td>
<td>41.9%</td>
<td>14.5%</td>
<td>85.5%</td>
<td>51.1%</td>
<td>48.9%</td>
</tr>
<tr>
<td>Valentine, Pickering, &amp; Darling (2003)</td>
<td>584</td>
<td>237</td>
<td>121</td>
<td>226</td>
<td>40.6%</td>
<td>20.7%</td>
<td>38.7%</td>
<td>61.3%</td>
<td>66.2%</td>
<td>33.8%</td>
</tr>
<tr>
<td>Wells, Steblay, &amp; Dysart (2014)</td>
<td>494</td>
<td>132</td>
<td>75</td>
<td>287</td>
<td>26.7%</td>
<td>15.2%</td>
<td>58.1%</td>
<td>41.9%</td>
<td>63.8%</td>
<td>36.2%</td>
</tr>
<tr>
<td>Wixted, et al. (2016)</td>
<td>348</td>
<td>114</td>
<td>104</td>
<td>130</td>
<td>32.8%</td>
<td>29.9%</td>
<td>37.4%</td>
<td>52.6%</td>
<td>47.7%</td>
<td>52.3%</td>
</tr>
<tr>
<td>Wright &amp; McDaid (1996)</td>
<td>1561</td>
<td>611</td>
<td>310</td>
<td>640</td>
<td>39.1%</td>
<td>19.9%</td>
<td>41.0%</td>
<td>59.0%</td>
<td>66.3%</td>
<td>33.7%</td>
</tr>
<tr>
<td>Wright &amp; Skagerburg (2007)</td>
<td>134</td>
<td>78</td>
<td>28</td>
<td>28</td>
<td>58.2%</td>
<td>20.9%</td>
<td>20.9%</td>
<td>79.1%</td>
<td>73.6%</td>
<td>26.4%</td>
</tr>
<tr>
<td>Overall Sum</td>
<td>6734</td>
<td>2746</td>
<td>1599</td>
<td>2389</td>
<td>40.8%</td>
<td>23.7%</td>
<td>35.5%</td>
<td>64.5%</td>
<td>63.2%</td>
<td>36.8%</td>
</tr>
<tr>
<td>Weighted means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>40.8%</td>
<td>23.7%</td>
<td>35.5%</td>
<td>64.5%</td>
<td>63.2%</td>
<td>36.8%</td>
</tr>
</tbody>
</table>