Eyewitness Identification: Probative Value, Criterion Shifts, and Policy Regarding the Sequential Lineup
Gary L. Wells

Current Directions in Psychological Science 2014 23: 11
DOI: 10.1177/0963721413504781

The online version of this article can be found at:
http://cdp.sagepub.com/content/23/1/11
Eyewitness Identification: Probative Value, Criterion Shifts, and Policy Regarding the Sequential Lineup

Gary L. Wells
Iowa State University

Abstract
Approximately 75% of DNA exonerations are cases involving mistaken identification. Lab-based experiments by psychological scientists have informed the legal system about ways to reduce the misidentification problem. One of these ideas, the sequential lineup (which shows the witness one lineup member at a time), increases the ratio of accurate to mistaken identifications compared with the traditional simultaneous lineup (which shows the witness all lineup members at once). Gronlund, Wixted, and Mickes (2014, this issue) noted that the improvement in this ratio from the sequential procedure is the result of a conservative criterion shift rather than of an increase in discriminability. Although data support that interpretation, the data do not negate the fact that probative value is higher for the sequential lineup. The question for policymakers is whether a more conservative decision criterion is desirable. Considerations include the natural asymmetry between the errors of mistaken identification versus nonidentification and the relevance of an accumulated body of archival data that suggest that witnesses in actual cases are using loose decision criteria for making identifications.

Keywords
eyewitness identification, sequential lineups, probative value, decision criterion, mistaken identification

Eyewitness-identification evidence obtained from lineup procedures was never subjected to empirical tests of reliability until psychologists began conducting lineup experiments in the mid-to-late 1970s. It was not until the mid-1990s that the U.S. legal system began to take the issue seriously. A turning point was the advent of forensic DNA testing. Approximately 75% of the innocent people who have been exonerated by DNA testing have been associated with cases that involved mistaken eyewitness identification (Innocence Project, 2013). These exoneration cases include people such as Kirk Bloodsworth, a man who had never been in trouble with the law but who was convicted of murder on the basis of mistaken eyewitness identification and sentenced to Maryland’s death chamber. He was eventually exonerated by DNA testing, and the actual murderer has now been determined (Junkin, 2004). These exonerated people are the lucky ones because few crimes leave behind DNA-rich biological evidence.

The legal system’s newfound concern with the eyewitness-identification problem relates closely to the message that psychological scientists have been delivering for nearly 40 years, namely, that the tendency to believe eyewitness-identification evidence tends to exceed its probative value. Probative value refers to the strength of the relationship between the proffered evidence at trial and the proposition sought. Probative value is a revered term in the legal system and represents the extent to which a piece of evidence makes it more probable that a proposition (e.g., that this is the person who committed the offense) is true. For eyewitness-identification evidence, probative value is reflected in likelihood ratios or probabilities that an identification of the defendant offered at trial was accurate or mistaken. In eyewitness-identification experiments, the ratio of accurate to mistaken identifications obtained under some set of conditions, typically referred to as diagnosticity, is considered an index of the probative value of an identification (Wells & Lindsay, 1980).

Corresponding Author:
Gary L. Wells, Department of Psychology, Iowa State University, West 112 Lago Hall, Ames, IA 50011
E-mail: glwells@iastate.edu
The legal system's interest in improving the probative value of eyewitness-identification evidence dovetails with eyewitness researchers' focus on system variables, which are variables that affect the accuracy of eyewitness-identification evidence over which the justice system has control (Wells, 1978). The research presented here addresses some of the issues surrounding one system variable, namely, simultaneous versus sequential lineup procedures.

A simultaneous lineup is the traditional type of lineup in which the witness views all lineup members at once. A sequential lineup is one in which the witness views lineup members one at a time without knowing the number to be viewed. Meta-analyses have shown that the probative value of the sequential procedure is higher than the probative value of the simultaneous procedure (Clark, 2012; Steblay, Dysart, & Wells, 2011). Gronlund et al., 2014, this issue) argued that this result is due to the fact that the sequential procedure is a more conservative test (an upward shift in the decision criterion) rather than that the sequential procedure is an improvement in discriminability over the simultaneous procedure. This is not a new point about the sequential procedure (e.g., see Meissner, Tredoux, Parker, & MacLin, 2005; Palmer & Brewer, 2012), and I agree that the empirical evidence is consistent with this argument. Palmer and Brewer (2012), for example, analyzed 22 experiments and concluded that although the sequential procedure does not increase discrimination, it does promote a less-biased criterion setting.

My aim in the present research was to place the Gronlund et al. (2014) article in a broader context, counter some questionable claims, and consider the policy implications. First, the sequential-superiority argument in the literature has always been made on the basis of an assessment of its probative value for identifications. Even if this better probative value comes from a criterion shift, an identification obtained from a sequential lineup can be better trusted than can one from a simultaneous lineup. Of course, if the increase in the probative value of identifications is due to an increase in the decision criterion (rather than to a change in discrimination), then there will be a decrease in identifications of the culprit along with a decrease in identifications of the innocent. This pattern was observed in simultaneous versus sequential lineup comparisons from the outset (Lindsay & Wells, 1985), and the pattern holds in meta-analyses (Steblay, Dysart, Fulero, & Lindsay, 2001; Steblay et al., 2011).

The simple way to think about this distinction is that eyewitnesses are less likely to make an identification with the sequential procedure than with the simultaneous procedure, but when they do make an identification with a sequential procedure, it is more trustworthy for the prosecutor, the judge, and the jury. At the same time, of course, the more conservative sequential procedure will produce more “misses” (failure to identify the culprit) than will the simultaneous procedure. But it should be kept in mind that a miss does not necessarily mean that a culprit will go free. People are convicted in great numbers without the aid of eyewitness-identification evidence because guilty people tend to have other evidence against them even if there is no eyewitness. Nevertheless, this distinction between the two procedures inevitably raises policy discussions about the relative weights that ought to be attached to the errors of identifying the innocent versus not identifying the guilty.

The Probative Advantage of the Sequential Lineup

Gronlund et al. (2014) occasionally questioned whether there is a probative value (diagnosticity) advantage for the sequential procedure, but they did so by pointing to a few selected contrasts in a few studies rather than by relying on broad meta-analyses. All literatures have some outliers. The better probative value of the sequential lineup is well documented by meta-analyses and reviews of the extant literature. Steblay et al. (2001) and Steblay et al. (2011) meta-analyzed 72 tests of simultaneous versus sequential lineups from 23 labs involving more than 13,000 participant witnesses. Across these studies, the accurate-to-incorrect identification ratio (probative value) for the sequential procedure was 6.3, whereas the ratio for the simultaneous procedure was 5.5. In a subanalysis that included only the 27 studies in which researchers employed the full design and did not have ceiling or floor effect problems, the sequential ratio was 7.7 and the simultaneous ratio was 5.8. Clark (2012) used slightly different methods and a slightly different set of studies and reported accurate and inaccurate identification rates across 51 tests of sequential and simultaneous procedures that correspond to ratios of 4.8 and 3.6, respectively.

Palmer and Brewer (2012) analyzed 22 simultaneous versus sequential studies and, as with the other meta-analyses, the ratio of accurate to mistaken identifications (probative value) was greater for the sequential procedure. But Palmer and Brewer also calculated both d' (a measure of discrimination) and c (a measure of response bias or decision criterion). There was no significant difference in discrimination, but there was a significant difference in response bias. A c score of 0 indicates unbiased responding, and negative values indicate a bias to identify. “Although responding was markedly lenient for both presentation modes, witnesses—on average—set a less-biased response criterion for sequential lineups [c = −.41]
than for simultaneous lineups \[ c = -0.94 \] (Palmer & Brewer, 2012, p. 252). In other words, witnesses were too eager to make identifications with both procedures but less so with the sequential lineup.

**Is There Evidence That Eyewitness Decision Criteria Are Too Low in Actual Cases? Archival Data**

Should the decision criterion be raised for eyewitness-identification evidence? That is, of course, a policy issue. Relevant to this question, however, are eight published archival analyses of actual lineups conducted in various locations in the United States and the United Kingdom. Because a proper lineup contains known-innocent fillers who are included in the lineup for the purpose of maintaining a procedure that is fair to the suspect, researchers are able to count how often witnesses do not identify anyone and how often witnesses identify known-innocent fillers. For example, in an archival study in the United Kingdom, researchers analyzed more than 1,000 lineups and found that 65% of witnesses made an identification of someone and that among those witnesses who made an identification, 40% identified an innocent filler (Horry, Memon, Wright, & Milne, 2012).

Table 1 lists and provides identification outcomes for all eight published studies in which researchers have examined how often witnesses in serious criminal cases identify a known-innocent filler. Overall, one of every three witnesses who made an identification chose a filler, even though the witnesses were clearly instructed that the culprit might not be present and, thus, the witnesses could (and should) have made no identification. Although fillers will not be charged with a crime, any one of those fillers could have been an innocent suspect. Moreover, it is unknown how often the suspect in those lineups was innocent; thus, the filler identification rate is an underestimate of the actual error rate among those witnesses who made an identification.

Given that at least one third of actual eyewitnesses to serious crimes who make an identification are choosing a known-innocent lineup member, a reasonable person could conclude that eyewitnesses’ criteria for choosing are too lax and that a more conservative procedure is an appropriate policy response to this problem. Receiver operating characteristic curves are a nice statistical tool, but they cannot answer an important question that policymakers face in this situation: Should we use a more conservative procedure that will improve the ratio of accurate to mistaken identifications?

### Identifications and Nonidentifications: The Policy Issue

Any suggestion that the ratio of accurate to mistaken identifications is not better with the sequential procedure seems unfounded given the meta-analyses that have shown otherwise. Gronlund et al. (2014) are on safer grounds, however, when they point out that a better ratio of hits to mistaken identifications resulting from the sequential procedure appears to be due to a criterion shift rather than to a change in discrimination. As a result, improvement in the ratio of accurate to mistaken identifications will be accompanied by a poorer ratio of correct to false rejections. How have policymakers dealt with this issue?

Many U.S. jurisdictions now use sequential lineup procedures, including the states of New Jersey, Connecticut, and North Carolina, as well as individual cities and counties, such as Dallas, Boston, Denver, and Santa Clara County, California. During the process of deciding whether to adopt the sequential procedure, were these

<table>
<thead>
<tr>
<th>Study</th>
<th>Percentage of witnesses making an ID</th>
<th>Percentage of IDs that are of the suspect</th>
<th>Percentage of IDs that are of a known-innocent filler</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behrman and Davey (2001)</td>
<td>74</td>
<td>68</td>
<td>32</td>
</tr>
<tr>
<td>Behrman and Richards (2005)</td>
<td>67</td>
<td>78</td>
<td>22</td>
</tr>
<tr>
<td>Horry, Halford, Brewer, Milne, and Bull</td>
<td>64</td>
<td>72</td>
<td>28</td>
</tr>
<tr>
<td>(in press)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horry, Memon, Wright, and Milne (2012)</td>
<td>65</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>Memon, Havard, Clifford, Gabbert, and Watt (2011)</td>
<td>86</td>
<td>51</td>
<td>49</td>
</tr>
<tr>
<td>Valentine, Pickering, and Darling (2003)</td>
<td>62</td>
<td>66</td>
<td>34</td>
</tr>
<tr>
<td>Wright and McDaid (1996)</td>
<td>59</td>
<td>66</td>
<td>34</td>
</tr>
<tr>
<td>Wright and Skagerberg (2007)</td>
<td>79</td>
<td>73</td>
<td>27</td>
</tr>
<tr>
<td>Average across studies</td>
<td>70</td>
<td>67</td>
<td>33</td>
</tr>
</tbody>
</table>
jurisdictions misunderstanding what the data showed? Definitely not. Having worked with most of these jurisdictions, I know that the primary data they used at that time was the Steblay et al. (2001) meta-analysis, which showed very clearly the trade-off between accurate and mistaken identification rates inherent to the sequential procedure. Far from being cavalier about the simultaneous versus sequential issue, the policymakers engaged in deep discussions and displayed considerable sophistication in their thinking.

**The status quo effect**

Policymakers are quite aware of the status quo effect. Accordingly, when they ask whether they should adopt a new procedure (sequential) that reduces mistaken identifications but increases the chances that the guilty might not be identified, they always try to “flip the status quo.” Specifically, they also ask, What if we had always used a sequential procedure and then someone came along and said we should use a simultaneous procedure? The question then becomes, Should we adopt a new procedure (simultaneous) that increases the chances that the guilty might be identified but also increases the chances of mistaken identification?

**The mistaken identification/nonidentification asymmetry**

Although value judgments come into play, policymakers are well aware of a natural asymmetry between not identifying the guilty and mistaken identifications of the innocent that is not in and of itself a value judgment. Specifically, a failure to identify the guilty party creates one potential error, namely, that a guilty person might go free. But the identification of an innocent suspect creates two potential errors: an innocent person could be convicted and a guilty person might go free. This asymmetry is a very interesting one because it is not a value judgment that places greater weight on one type of error as opposed to another. Instead, it is a judgment that one error (a false rejection) cannot be as serious as that same error (a false alarm) plus another error (a false alarm).

**The second-lap policy**

Because they recognize that the sequential procedure could result in an eyewitness failing to identify the culprit, every U.S. jurisdiction that has adopted the sequential procedure has included a proviso that the witness can view the sequence a second time if the witness explicitly requests to see the lineup members again. Of course, any second lap through the sequential lineup is a matter of record that must be disclosed to the defense, and it could be argued that a second lap regresses to the functional equivalent of a simultaneous procedure for that witness. However, this compromise is interesting because it separates witnesses who were able to make their identification before knowing what the other lineup members looked like from those witnesses who felt that they had to see the remaining lineup members to make a decision. The proviso is similar to the observation made by Steblay et al. (2011) that you can always perform a sequential procedure followed by a simultaneous procedure but it makes no sense to do the reverse.

**Additional policy considerations**

Policymakers look well beyond the data to include such things as the cost of lawsuits associated with wrongful conviction, the need for greater probative value in evidence, and the cost that convicting the innocent has on the public’s confidence in the justice system.

**Final Remarks**

Gronlund et al. (2014) were correct to suggest that receiver operating characteristic analyses are the best way to determine whether the simultaneous/sequential difference is a criterion shift. But equality of receiver operating characteristic curves does not negate the fact that probative value is greater for the sequential procedure. When it comes to considering lineup reforms, the primary concern of policymakers has been that witnesses too often make mistaken identifications, not that they too often fail to identify the culprit. The sequential lineup is one tool that some jurisdictions have decided is a useful one for the goals they have in mind.

It is important to note that the argument that the sequential procedure increases the decision criterion is not an explanation but instead is a mere redescription of the data. Psychologically, it remains quite viable to argue that the sequential procedure is more conservative (a higher criterion) precisely because witnesses have to make an identification decision before knowing what the other lineup members look like (a more absolute decision rather than a relative decision); that is, the sequential procedure prevents witnesses from using a lax criterion of simply deciding which lineup member, compared with the other lineup members, looks most like the culprit (which would lead to a mistake whenever the culprit is not in the lineup). Within the sequential procedure, the witnesses know that there might be someone later in the sequence who looks even more like the culprit. What should one make of a witness who got to the fourth lineup member in a sequential procedure and said, “Can I see the rest before I decide whether this is the person I want to identify as the murderer?” Given the powerfully
incriminating nature of eyewitness-identification evidence, it seems that eyewitnesses ought to be able to show that they can identify the culprit if they see the person without needing to see the remaining lineup members.

I agree with Gronlund et al. (2014) that the goal should be to find even better eyewitness-identification procedures (see Brewer & Wells, 2011). If the traditional lineup did not already exist, a wholly different approach probably would be developed, perhaps one involving eye movements, pupil dilation, event-related potential patterns, response latencies, implicit memory tests, and other potential indicia of recognition. Bringing psychological science to bear on the serious problem of eyewitness identification ought to mean much more than manipulating whether photos are shown as groups versus one at a time. The next generation of eyewitness researchers should throw out the traditional lineup approach and bring more creative tools to bear.

Recommended Reading


Author Contributions

G. L. Wells is the sole author of this article and is responsible for its content.

Declaration of Conflicting Interests

The author declared that he had no conflicts of interest with respect to his authorship or the publication of this article.

Note

1. A recent national survey estimated that 32% of U.S. law enforcement agencies use the sequential procedure (see Police Executive Research Forum, 2013).

References


Steblay, N. K., Dysart, J. E., & Wells, G. L. (2011). Seventy-two tests of the sequential lineup superiority effect: A