Eyewitness identification research has had a strong impact in the United States recently. A. M. Levi and R. C. L. Lindsay (2001) discussed the role of psychologists’ recommendations for lineup reform and urged that these include numerous items, however complex, rather than only a few items that are readily understood by policymakers. They focused on the recommendations of 1 article in particular (G. L. Wells et al., 1998) and argued that this article should have called for abolishing show-ups, using more lineup fillers, adding voice to lineups, and permitting witnesses to make multiple choices. In the present article, the author raises concerns about each of these 4 recommendations and defends the idea of researchers culling their recommendations to a focused set. Finally, the author describes the emerging positive impact of the Wells et al. recommendations on the reform of lineup procedures in the United States.

Prompted by new insights into the psychology of eyewitnesses to crimes, New Jersey is changing the way it uses witnesses to identify suspects. Starting in October, the state will become the first in the nation to give up the familiar books of mug shots and to adopt a simple new technique called a sequential photo lineup, said John J. Farmer Jr., New Jersey’s attorney general.


The new rules for police lineups in New Jersey represent a significant advance in the impact of eyewitness research in the United States. New Jersey not only changed to the sequential procedure, which shows only one photo at a time and requires a decision before seeing the next photo, but also adopted double-blind testing, a procedure in which the person who administers the lineup does not know which person is the suspect in the case. These are procedures that eyewitness researchers have been pressing for in recent years. Not only is this a first for the United States, but it appears that there is no other large jurisdiction in the world that has built both of these safeguards into its lineup procedures. Although some parts of Ontario, Canada, use sequential procedures, double-blind testing is not included. Prosecutors and police in other jurisdictions across the United States are now considering similar changes to those in New Jersey.

The current developments in the United States contrast dramatically with the situation only a few years ago when the legal system actively ignored the eyewitness literature (except when forced to deal with the narrow issue of the admissibility of expert testimony). The path to this newfound impact of eyewitness research on police practices is traceable in part to two sets of guidelines, the American Psychology–Law Society (AP-LS) guidelines, hereinafter called the Lineups White Paper (Wells et al., 1998), and the National Institute of Justice’s

Correspondence concerning this article should be addressed to Gary L. Wells, Department of Psychology, Iowa State University, Ames, Iowa 50011. Electronic mail may be sent to glwells@iastate.edu.
One purpose of the current article is to address an article by Levi and Lindsay (2001) that argued that the Lineups White Paper should have included a number of additional recommendations and that recommendations papers by psychologists should generally include all possible improvements based on research findings. The Lineups White Paper clearly did not take this approach. Instead, the Lineups White Paper focused primarily on four recommendations. I continue to believe that the focused approach is the right one and that the impact of the Lineups White Paper is attributable at least in part to this focused approach. A consideration of my arguments for a focused approach and of Levi and Lindsay’s arguments for a broad approach could help other researchers make this type of judgment for themselves in other domains of the interface of psychology and policy.

A second purpose of this article is to discuss some of the specific recommendations that Levi and Lindsay (2001) argued should have been included in the Lineups White Paper. I use their examples to illustrate the distinction between the relatively narrow focus of experimental data and the broader focus of legal policy considerations. In some cases, such as the discussion of one-person lineups (called show-ups), this is the first published discussion of the concerns that policymakers have and how those concerns are not directly addressed by experimental data.

Relation Between the Lineups White Paper and the NIJ Guide

A brief overview is useful to understand how the Lineups White Paper is related to the NIJ Guide. The Lineups White Paper was a document solicited by the executive committee of the AP-LS to recommend improvements to how lineups are conducted based on scientific studies for which there would be broad agreement by scientists. It was designed to be highly readable so that individuals outside of the science, such as prosecutors and police, would be able to read and understand the material. Although many have focused on the recommendations that were contained in the Lineups White Paper, perhaps the most important feature of the Lineups White Paper was the inclusion of an analysis of the first 40 DNA exoneration cases, which showed that mistaken eyewitness identification was the primary evidence used to convict these innocent people in 36 of the 40 cases. The Lineups White Paper was the first refereed publication that examined DNA exoneration cases to show that mistaken identification was at the heart of these convictions of innocent people.

A prepublication draft of the Lineups White Paper was read by then-U.S. Attorney General Janet Reno. Reno was impressed with two elements of the Lineups White Paper. First, the Lineups White Paper’s analysis of the DNA exoneration cases confirmed and extended an earlier study that she had commissioned of the first 28 DNA exoneration cases. Second, she was impressed with the general argument that there exists a body of research showing how improvements to lineup procedures could help reduce the mistaken identification problem. Reno then ordered the National Institute of Justice (NIJ; the research and education arm of the U.S. Department of Justice) to form a panel to make recommendations for improving eyewitness evidence. Reno made it clear that eyewitness researchers,
rather than just police and lawyers, were to be central members of the panel. The written product that resulted from this panel and the working group that followed it was the NIJ Guide.

A more complete version of the process that produced the NIJ Guide can be read elsewhere (Wells et al., 2000). It is important to note, however, that the Lineups White Paper was certainly not the only source used by the technical working group that developed the NIJ Guide. In fact, with five eyewitness researchers on the working group that wrote the NIJ Guide, the larger working group was bombarded with information and findings that spanned a broad range of eyewitness issues not covered in the Lineups White Paper. Although the Lineups White Paper was prominent in the deliberations of the NIJ working group, it was one of many sources of information.

The Focused Approach of the Lineups White Paper

The idea of focusing on a small number of recommendations was based on several considerations as described in the Lineups White Paper. First, there have been many previous publications that have contained large numbers of recommendations on improving lineups, in some cases containing over 100 recommendations (Wells, 1988). Inevitably, this dilutes the more important, central recommendations and yields a product that more closely resembles a literature review than a policy recommendation. A small, easy to remember, focused set, on the other hand, was thought to have a greater potential for having an impact on policymakers. Levi and Lindsay (2001) made some good arguments about the idea of including all possible recommendations in such policy arguments, but I continue to believe that the focused approach is superior, especially under conditions in which policymakers have been ignoring, and even actively counterarguing, the need for changes.

Clearly, any long list of recommendations is more likely to have weak, unwise, poorly thought out, or even unreliable recommendations than is a shorter, more carefully culled list. Researchers who venture into the policy domain must be careful not to mix their stronger with their weaker recommendations. Policymakers who are unfriendly to the general idea of change will inevitably focus on the weakest elements in the recommendations to discredit the source. Indeed, the researchers who worked on the NIJ Guide were constantly under attack, primarily from prosecutors on the working group, for being unrealistic about police work or proposing ideas that were not well thought out. Had the Lineups White Paper included some weak or impractical recommendations, Reno might have never formed the panel in the first place, or she might not have insisted that researchers be a part of the panel, or researchers might have had less credibility in the working group. With the Lineups White Paper we had a foot in the door with four focused, difficult-to-counterargue recommendations.

Should the authors of the Lineups White Paper have included more recommendations? Perhaps they should have. As explained in the Lineups White Paper itself, one recommendation, a recommendation that simultaneous lineups be phased out in favor of sequential lineups, was strongly considered. Had more than four recommendations been made, the sequential lineup would have been included. However, a principal reason for not including the sequential lineup
procedure is clearly stated in the Lineups White Paper. In particular, our goal was to make recommendations that could stand independently of each other. In other words, if policymakers were to adopt any one of these four recommendations, then that would be a desirable outcome. We believed, and there is now empirical evidence to confirm our belief, that this is not true of the sequential procedure. Specifically, the sequential procedure may actually be worse than the simultaneous procedure under conditions that do not use double-blind testing or some other method to prevent the influence of the person administering the lineup (Phillips, McAuliff, Kovera, & Cutler, 1999). Sequential lineups without double-blind testing may be worse than simultaneous lineups without blind testing. Hence, the sequential lineup recommendation does not pass the test of independence from the other recommendations. That argument was clearly made in the Lineups White Paper.

Of course, the authors of the Lineups White Paper could have dropped the requirement that the recommendations meet the independence test. However, these are the types of decisions that researchers must make at the time of making recommendations, knowing full well, of course, that others might ultimately second-guess them years later.

Although I might agree that the sequential procedure should have been included so that there were five core recommendations in the Lineups White Paper, other recommendations, including some that Levi and Lindsay (2001) argued should have been included, seem to me, even in hindsight, to be unwise. In the following sections, I address four recommendations mentioned by Levi and Lindsay that help illustrate some basic distinctions between laboratory data and legal policy.

**Broad Policy Considerations Versus Isolated Considerations of Data**

Although psychologists are well trained in research, few have ever taken a course in policy studies or read policy journals that would give them a perspective on the multifarious issues involved in the development of public and legal policy. Policy considerations regarding lineup procedures are complex, and most of these complexities are not reflected or contained in the laboratory data of the eyewitness researcher. Levi and Lindsay (2001) alluded to this distinction, but it is much more complex and important than their analysis suggests. One such policy concern that prosecutors face, for instance, is that the formal declaration of new, improved lineup procedures could lead to appeals by tens of thousands of previously convicted persons based on the argument that they were convicted using the inferior lineup procedures. This could, in turn, completely stifle the already overloaded postconviction appeals system, and as a result, truly meritorious appeal cases could be hidden in a sea of frivolous cases. It might be easy for researchers to dismiss this argument in their own minds (e.g., by saying that these people should be allowed such appeals), but that is my point: Researchers do not have to worry about such consequences whereas policymakers do. Eyewitness researchers who summarily dismiss such concerns have no credibility with policymakers, and those without credibility with policymakers have no chance to shape policy.

In the following sections, I address Levi and Lindsay’s (2001) arguments for
banning show-ups (single-person lineups), permitting witnesses to make more than one choice, using a large number of fillers for lineups, and adding voice to lineups. In each case, I note policy concerns.

Levi and Lindsay (2001) Called for a Ban of Show-Ups

There is perhaps no clearer example of the distinction between legal policy and experimental data than in the area of show-ups. A show-up is an identification procedure in which police present a single suspect to the eyewitness(es) to see if the eyewitness(es) will identify that person as the perpetrator. Unlike a lineup, there are no fillers (distractors, foils) in a show-up procedure. Show-ups are legal in the United States if the crime occurred recently (within the past hour or 2 at most) and a person who fits the description is detained in the area. Experimental data clearly show that show-ups are inferior to lineups in terms of the chances of a mistaken identification (e.g., Dekle, Beale, Elliot, & Huneycutt, 1996; Gonzalez, Ellsworth, & Pembroke, 1993; Lindsay, Pozzulo, Craig, Lee, & Corber, 1997; Yarmey, Yarmey, & Yarmey, 1996).

Before going further with this discussion of show-ups, it is important to understand what the data comparing show-ups to lineups really look like. It is commonly said of a show-up that it is a highly suggestive procedure compared with a lineup. Whatever might be meant by the term highly suggestive, it should be noted that the data do not support the idea that eyewitnesses are more willing to make an identification with a show-up compared with a lineup. In fact, the choosing rates with show-ups appear to be lower than the choosing rates with lineups. This commonly overlooked pattern holds both for real case data in the field (Gonzalez et al., 1993) and for laboratory data (e.g., Dekle et al., 1996; Lindsay et al., 1997; Yarmey et al., 1996). Hence, the superiority of the lineup over the show-up is not because the show-up makes witnesses more willing to identify someone. Instead, the lineup is superior because it can control errors by spreading errors to the fillers (“known errors”), whereas an error with a show-up is always an error of mistakenly identifying a suspect. This pattern in the data must be kept in mind when evaluating the legal system’s analysis of show-ups.

Legal policy necessarily takes a broad view of the issue of show-ups. Among the broader issues for legal policymakers are the need to free an innocent person from suspicion as quickly as possible and the need for public safety.

Consider first the need for a quick release of innocent persons from suspicion. Show-ups detain the suspect for only a matter of minutes and are not considered arrests. Field data indicate that 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”), thereby permitting a quick exoneration of the detained person with minimal disruption to his or her life (Gonzalez et al., 1993). What is the alternative to permitting the show-up? To have the time to conduct a full lineup procedure, complete with appropriate fillers, the person detained on the street would have to be arrested. Should we arrest every suspect rather than permit a show-up? This would be unconstitutional in the United States because there would not be sufficient probable cause to make an arrest. Even if it were legally permissible to arrest the suspect, is this sensible or
prudent in light of the assumption that many, if not most, of these people are in fact innocent?

There is also the issue of public safety. If police cannot conduct a show-up because show-ups have been banned and police cannot arrest the suspect because it would violate the suspect’s constitutional rights, what should they do? Should the detained person routinely be set free, knowing that those who are guilty will flee and perhaps be a danger for others? Public safety is a legitimate concern for policymakers.

Reasonable people can disagree about how to answer these questions, but the simple observation that show-ups are not as effective as lineups for purposes of identification evidence is not the only consideration in deciding whether or not to permit show-ups as a matter of legal policy. Law enforcement and other participants and stakeholders in the legal system have a number of considerations to balance on this issue. It is instructive to note that none of the five published studies contrasting lineups with show-ups has ever discussed the policy conflict that I have described here. This reinforces my point about the absence of deliberated policy considerations among psychological researchers.

None of this excuses the legal system for failing to act on this important problem, but it highlights the distinction between a narrow focus on laboratory data and the larger considerations that are required by policymakers. It might or might not be advisable to ban show-ups, depending on how one resolves policy considerations about public safety and the need for a quick release of innocent people from suspicion rather than arresting them. Had the Lineups White Paper maintained the hard-line position that show-ups should never be conducted, policymakers would have concluded, perhaps reasonably, that the researchers were incapable of appreciating the broader implications of their arguments. This, in turn, might have led to ostracism of the researchers by the police, who constituted the majority of the members of the working group that wrote the NIJ Guide.

The show-up presents a policy dilemma, pitting the suggestiveness of the procedure against the need to free innocent persons quickly and the need for public safety. In the long run, this policy dilemma might be solved with technology. Consider, for instance, police cruisers with cameras, dedicated photo-lineup software on laptop computers, and wireless Internet access to a large bank of filler photos. The witness on the street might then be able to view the suspect in a photo lineup with appropriate fillers via a laptop computer display, without requiring that the police arrest the suspect or detain the suspect for more than the few minutes that it currently takes to conduct a show-up. Researchers can serve the cause of justice more effectively by developing the software for field identification alternatives to the show-up than they can by simply calling for a ban on show-ups. Constitutional issues might nonetheless be implicated, because it is unclear whether police may photograph someone who is not charged with a crime under these circumstances. The courts would have to determine the answer to that question, again based on broad legal policy considerations.
Levi and Lindsay (2001) Recommended That Witnesses Be Allowed to Select More Than One Person

Among the recommendations that Levi and Lindsay (2001) suggested should have been included in the Lineups White Paper is that eyewitnesses be allowed to make more than one choice from a sequential lineup (even though there is only one perpetrator). This idea stems from a study by Levi (1998b).

Even if the Levi (1998b) article had been published in time to include it in the Lineups White Paper, there is no way that it would have been included among the recommendations. Perhaps there are some special characteristics of Israeli criminal law that will make this work feasibly in Israel, but prosecutors and defense attorneys in the United States would resist this idea, each for different reasons. To the defense attorney, this practice encourages choosing and thereby increases the chances that his or her client will be identified. To the prosecutor, selecting more than one person when there was only one perpetrator is tantamount to having an impeached witness before the witness even takes the stand.

I encourage Levi and Lindsay (2001) to continue to develop this multiple-choice idea, but at this point in its development I know of no researchers other than Levi and Lindsay who have argued for this change. This seems to me to be a clear example of the distinction between a review paper, which almost certainly includes mention of the Levi (1998b) multiple-choice idea, and a policy recommendation paper. In order to graduate this idea to the level of a policy recommendation, it will need to survive not only peer review and be broadly accepted by researchers, but also the careful scrutiny of policymakers who will examine it thoroughly for potential negative effects in some other domain.

Levi and Lindsay (2001) Recommended That Voice Be Added to the Lineup

Neither the NIJ Guide nor the Lineups White Paper discussed voice identification. Perhaps this was an aspect that should have been included, but the guidelines were focused on visual identification, not voice identification. More important, this is an issue that must be more deeply thought out and researched. Although there are now quite a bit of data on voice identification from voice lineups, there are comparably little data on using voice and visual identification together in a live lineup. Although there is a study indicating a benefit for voice plus visual lineups together (Melara, DeWitt-Rickards, & O’Brien, 1989), there are dynamic, real-world problems not addressed in that work. First, in a live lineup, the fillers are selected based on their physical appearance, not their voices. Hence, a lineup might be “fair” based on the physical appearance of the fillers but biased based on the voices. Suppose, for example, someone became a suspect in part because she or he had a particular accent. A simple recommendation that lineup members should speak could be an unwise recommendation. Voice lineups require that the filler voices meet their own set of criteria, and the usual ways of selecting lineup fillers based on visual information would not satisfy these criteria. A second problem with using voice and visual lineups together is that it can create ambiguity as to what aspect the eyewitness used in making the decision and in what sequence. If the eyewitness saw the lineup, then heard the voices, then made
an identification decision, was the decision responsive to the voice, the visual information, or both? One can ask the witness whether the decision was based on the voice, the physical appearance, or both, of course, but there is a risk in trusting the eyewitness’s self-reports.

Perhaps it would be best, therefore, to construct a separate voice lineup and never permit the lineup members to speak at the time of the visual lineup. The voice lineup would probably not include the same fillers as the visual lineup. This separate procedure not only takes care of the problem of selecting a fair set of fillers for the voice but also makes it clear what the witness is identifying. There is the added bonus here that selection of the suspect in both the visual and voice lineups would reduce the possibility even further that mere chance is involved because the probabilities can be multiplied according to independence rules for the two lineups.

On the other hand, it is not yet known whether separating the voice and visual lineups is the best procedure because there are no published data to answer this question. As we learn more about the optimal ways to conduct voice and visual lineups, in particular contrasting the joint versus separate procedures, we can begin to incorporate these recommendations into formal proposals. However, until we know the answer based on peer-reviewed publications and replication, we should not pretend to know the answer. Again, a greater certainty is required to include something in policy recommendations than is required to include something in a literature review. It is easy to cite a study suggesting that voice and visual information together seem to improve eyewitness performance; it is another thing to make the transition from this study to a clear policy recommendation.

**Levi and Lindsay (2001) Recommended More Fillers**

The easiest recommendation of Levi and Lindsay’s (2001) to agree with is that the number of fillers used currently by law enforcement should be increased. Levi and Lindsay were correct in the assertion that the chances of mistaken identification of a suspect will decrease as the number of fillers who fit the description is increased. Levi and Lindsay seemed particularly incensed that both the Lineups White Paper and the NIJ Guide mention the number six for lineup size using photographs and five with live lineups. However, these lineup sizes are clearly specified in plain language in both sets of guidelines as a minimum number. What Levi and Lindsay did not realize, perhaps, is that many lineups in the United States have consisted of only two or three fillers. Furthermore, it is well documented that the number of persons per se in a lineup is not nearly as important as the number of lineup members who fit the description of the culprit. The Lineups White Paper and NIJ Guide stressed the qualities of the fillers more than the quantity of the fillers.

Why not specify a much larger minimum of fillers? In many cases, it is nearly impossible to find more than five fillers who fit a description if the description has specific characteristics that are rare in the population (e.g., scars, minority racial status, large nose, protruding ears). A policy that sets a minimum lineup size of 20, as mentioned in the Levi and Lindsay (2001) article, would simply make it impossible to conduct a lineup at all in many cases. Again, one must think about the consequences of such a policy. It could be argued, for instance, that the
20-minimum lineup size policy would give perpetrators with unusual physical characteristics carte blanche to offend, with guarantees that they could never be placed in a legal lineup because that many fillers could not be found who fairly fit the description. Undoubtedly, systems could be created for police departments to share large banks of photos electronically, but this system is not currently in place and, until it is, large minimum number requirements for lineup sizes will not be accepted by law enforcement agencies.

Other Concerns About the Levi and Lindsay (2001) Approach

Although I have concerns about some of the items that Levi and Lindsay (2001) argued should have been included in the Lineups White Paper, it is not my intent to imply that the four recommendations in the Lineups White Paper were the only defensible ones to make at that point. I concede, for instance, that the sequential procedure could have been among these recommendations if the independence consideration had been dropped. It is my intent, however, to argue that it is better to make a small number of difficult-to-counterargue recommendations than it is to make a large number of recommendations that are not quite so well thought out. This is especially important in domains for which the policymakers are reluctant to make changes in the first place. Once policymakers are willing to make changes, then authors of recommendations publications can take a different approach involving a larger set of recommendations.

In general, I encourage researchers who are writing recommendations for policymakers to make a very sharp distinction between a review paper, which typically is intended to include everything relevant to the issue, and a policy recommendations paper, which would typically focus on a smaller set of ideas for which broader policy considerations have been well thought out. Furthermore, review papers are typically written for other researchers and, therefore, often include complex treatments of incomplete ideas written in a research jargon. Policy recommendations papers, on the other hand, must be simpler, more selective, and readable by nonpsychologists who are unfamiliar with the literature.

There is another element in the Levi and Lindsay (2001) article that I believe can be troublesome for purposes of a policy recommendations paper. In particular, they stated that “a perfectly fair standard simultaneous lineup may still result in 25% of identified suspects being innocent” (p. 787). This conclusion is based on a single article by Levi (1988a) that many researchers have not fully accepted, and I, for one, do not believe. Passing such statements to policymakers when they are not fully accepted by a broad range of researchers is unwise.

Furthermore, Levi and Lindsay (2001) stated that “staged crime experiments, using simultaneous lineup procedures comparable to those that would result from the four recommendations, routinely generate high rates of false-positive choices (averaging about 60%)” (p. 787). Statements such as these are misleading in my opinion. The 60% figure, for instance, includes all identifications in spite of the fact that five out of six of the people in a lineup are known-innocent fillers. In a fair six-person lineup for which the suspect is innocent, only one-sixth of the choices would be of an innocent suspect. If we accepted the 60% figure, then a more forensically valid estimate of the rate of innocent suspect identifications would be 10%. In fact, however, this 60% figure rests on a number of procedural
assumptions, including some questionable ways in which researchers select individuals to serve as the innocent suspect who stands in for the culprit in culprit-absent lineups. In addition, the 60% figure presumes that the lineup does not include the culprit. If the actual culprit is in the lineup and the single-suspect model is used as recommended, then there would be no mistaken identifications. Researchers still do not know how often lineups contain the actual perpetrator, how often they include an innocent suspect, and how often the innocent suspect happens to look a lot like the culprit (Wells, 1993). Hence, statements such as these do not belong in policy recommendations papers because they imply that we, as researchers, know more than we can in fact know. Furthermore, the presentation of such statements (e.g., 60% false-positive choices on average) is quite misleading and represents the type of statement that can harm our credibility with policymakers.

Final Comments

New Jersey stands as the prime example of the fact that neither the Lineups White Paper nor the NIJ Guide have served to limit the amount and type of improvements that policymakers can and will continue to make to lineups. Although the Lineups White Paper and the NIJ Guide were instrumental in influencing New Jersey to make improvements to lineup procedures, the New Jersey procedures are more ambitious and more complete than either the Lineups White Paper or the NIJ Guide. A focused, defensible set of recommendations can initiate a process that would otherwise have trouble catching hold.

I believe that it is healthy to challenge assumptions behind the strategy that was used in the Lineups White Paper, and the article by Levi and Lindsay (2001) represents a useful and thoughtful alternative perspective. I continue to believe that the Lineups White Paper took the correct approach by focusing on only four recommendations, but perhaps that view is to be expected from the senior author of that article. Researchers who step into the policy recommendations arena must not confuse a policy recommendations paper with a literature review paper. Literature review papers are inclusive; policy recommendations papers are selective. We will never know what would have happened if the Lineups White Paper had been written more like Levi and Lindsay think it should have been written. However, it is clear that the Lineups White Paper, as written, was a central piece of the puzzle that has now resulted in an immense and unprecedented move to improve eyewitness identification procedures in the United States. I believe that the defensibility, simplicity, and readability of the Lineups White Paper were critical to this impact.

References


