

2006

In Response to the Illinois Pilot Program on Simultaneous v. Sequential Lineups

Ebbe B. Ebbesen

Prof. of Psychology, University of California, San Diego

Kristin M. Finklea

Follow this and additional works at: <http://lawcommons.luc.edu/pilr>

 Part of the [Courts Commons](#), [Criminal Law Commons](#), and the [State and Local Government Law Commons](#)

Recommended Citation

Ebbe B. Ebbesen, & Kristin M. Finklea, *In Response to the Illinois Pilot Program on Simultaneous v. Sequential Lineups*, 11 Pub. Interest L. Rptr. 9 (2006).

Available at: <http://lawcommons.luc.edu/pilr/vol11/iss2/5>

This Article is brought to you for free and open access by LAW eCommons. It has been accepted for inclusion in Public Interest Law Reporter by an authorized administrator of LAW eCommons. For more information, please contact law-library@luc.edu.

SYMPOSIUM ISSUE



The Illinois Pilot Program data may show evidence of investigator influence and eyewitness bias, including racial bias.

In Response to the Illinois Pilot Program on Simultaneous v. Sequential Lineups¹

Ebbe B. Ebbesen² and Kristin M. Finklea

The true level of eyewitness accuracy in the legal system has been debated by researchers and laypersons alike. Specifically, inaccurate eyewitness identification is thought, by some, to be the primary cause of false convictions.³ The goal of the legal system is to maximize the number of convicted guilty suspects while minimizing (and, in theory, eliminating) the number of convicted innocent suspects. To minimize the rate of mistaken identifications, researchers have suggested that eyewitness identification evidence be collected using modified lineup procedures. Two of the major changes in protocol include: 1) using blind lineup administrators and 2) displaying the lineup photographs sequentially.⁴ In order to effect policy change in a scientifically reasonable manner, researchers should compare a proposed new policy against the established policy already in place. Without this comparison, we will never know whether the new policy is any better than the old. Such a comparison is no different than that used in the medical field to evaluate the effectiveness of a new treatment against the current method. The Illinois Pilot Program⁵ was designed utilizing this philosophy to compare the proposed sequential double-blind lineup procedure against the traditional simultaneous (non-blind) lineup procedure.

To date, findings from laboratory research suggest that witnesses make fewer selections from lineups

presented sequentially than they make from lineups presented simultaneously.⁶ As a consequence, both fillers and “suspects” are identified less often in sequential compared to simultaneous lineups. Based on these findings, some researchers have advocated that policy makers adopt sequential over simultaneous lineups to minimize the potential for eyewitnesses to mistakenly identify innocent suspects. Unfortunately, research has not adequately addressed whether adopting a sequential procedure in practice would also reduce the rate at which guilty suspects are identified.⁷ If the results of laboratory studies can be generalized to witness responses in actual criminal cases, both filler and suspect choice rates should decrease as a result of moving from a simultaneous to a sequential lineup procedure.

Summary of Major Findings from the Pilot Program

To assess these predictions, we can look at the results from the Illinois Pilot Program⁸ (Table 1). The entire sample contained a total of 367 different cases, in which researchers identified a total of 741 lineups. Of these, a total of 521 *unique* lineups were identified, as some investigators presented the same suspect in the same position with the same fillers to more than one witness. Across three jurisdictions investigators conducted a total of 366 standard simultaneous, single-suspect lineups and a total of 271 sequential, double-blind, single-suspect lineups. Witness/victims chose the suspect in 244 (or 67 percent) of all of the simultaneous lineups and in 154 (or 57 percent) of the sequential/blind lineups. Witness/victims chose fillers a total of 8 times (or 2.2 percent) when viewing the simultaneous lineups and 18 times (or 6.6 percent) when viewing sequential lineups (the difference in choice rates between the simultaneous and sequential lineup procedures was statistically significant). Including multiple suspect lineups in the analysis did not change the basic pattern of results. Thus, overall, the suspect choice rate was higher and the filler choice rate was lower for the simultaneous than sequential lineup procedure.

The tendency for witnesses to choose suspects more frequently and fillers less frequently given a

(Ebbesen/Finklea, *continued on page 10*)

(Ebbesen/Finklea, continued from page 9)

Table 1: Number and Percent of Suspect and Filler Choices for Known Single Suspect Lineups as a Function of Lineup Procedure

Lineup Procedure	Number of Choices			Percent of Total Choices		
	Suspects	Fillers	No Choice	Suspect	Filler	No Choice
Simultaneous	244	8	114	66.7	2.2	31.1
Sequential	154	18	99	56.8	6.6	36.5

simultaneous compared to a sequential lineup presentation was replicated for two of the three jurisdictions, Chicago and Evanston. In Chicago, suspects were chosen 64 percent of the time with simultaneous lineups and 49.5 percent of the time with sequential lineups. In addition, no fillers were chosen with simultaneous lineups, but with sequential lineups 6.3 percent of the choices were of fillers. Similarly, in Evanston, suspects were chosen 72.1 percent of the time with simultaneous lineups and 44.2 percent of the time with sequential lineups. Again, no fillers were chosen with simultaneous lineups, but 13.5 percent of the choices in sequential lineups were fillers. In Joliet, the pattern was slightly different. The suspect choice rates were 61.7 percent for simultaneous and 69.4 percent for sequential lineups, with filler choice rates at 4.4 percent and 6.5 percent, respectively. In summary, across all three jurisdictions, the known error rate (i.e. foil identifications) was higher with sequential than simultaneous lineups, and suspect choice rates were higher in simultaneous than sequential lineups for two out of three of the jurisdictions. It is clear that the results from this field program are in direct contradiction to the generalizations from laboratory research findings.

Alternative Explanations for the Findings

Two different classes of explanations might account for the differences in the pattern of results seen in the field project and that in laboratory simulations. We can focus on the differences in how lineups are constructed in laboratory research and how they are typically constructed in the legal system. Alternatively, we can focus on the fact that the two lineup procedures differed not only in terms of the presentation of lineup “alternatives” (simultaneously v. sequentially) but also

in terms of whether the investigators conducting the lineups knew the suspect’s location in the lineup.

Lineup Construction

The discrepancy between the Illinois Pilot Program⁹ results and the laboratory findings can be explained by considering that the characteristics of the lineups employed in the field might have differed from the characteristics of the lineups employed in the laboratory. First, the base rate of guilty suspects appearing in lineups might have been higher in the field compared to laboratory studies. In order to accurately assess the generalizability of laboratory research to eyewitness identifications in actual criminal cases, researchers should evaluate the true rate of target present (“TP”) and target absent (“TA”) lineups.¹⁰ In the typical laboratory study, there are equal proportions, 50:50, of TP and TA lineups. However, researchers have yet to determine if this proportion is representative of the rate at which guilty and innocent suspects appear in actual lineups. The burden of suspect choice errors is often overlooked in applied research, assuming that suspect equals culprit. The most problematic error an eyewitness can make is that of selecting an innocent suspect (not that of selecting a known-innocent foil) from a lineup. Some have concluded that in the laboratory, the largest difference in identification outcomes across the two lineup procedures occurs in the identification of “innocent” compared to “guilty” suspects.¹¹ That is, although both guilty and innocent suspects are chosen more often in simultaneous lineups, the difference found in innocent suspect identifications is larger.¹² Therefore, some have suggested that switching to a sequential lineup will reduce innocent suspect choices more than guilty

(Ebbesen/Finklea, continued on page 11)

SYMPOSIUM ISSUE*(Ebbesen/Finklea, continued from page 10)*

suspect choices. However, if guilty suspects are present in real world lineups more often than are innocent suspects, using a sequential lineup procedure will suppress the hit rate more than the false alarm rate in actual cases. As such, the higher rate of suspect choices in simultaneous compared to sequential lineups conducted in Illinois may have resulted from the fact that the large majority of suspects presented to witnesses for identification were actually guilty culprits rather than innocent suspects.

Second, the relatively low rate of filler choices in actual lineups compared to laboratory studies might be explained by the difference in the way fillers are selected for actual lineups compared to lineups constructed in a laboratory. In a typical controlled experiment, the guilty suspect is removed and replaced with an innocent look-alike. Foils, consequently, remain high in similarity to both the guilty (TP) and innocent (TA) suspects. In the real world, however, an innocent suspect who is apprehended may look nothing like the actual culprit. Hence, when foils are selected for the lineup based on their degree of similarity to the innocent suspect, these foils will have a high degree of similarity to the innocent suspect, but a low degree of similarity to the actual culprit. As a result, a witness may be less likely to select a foil from such a TA lineup than a lineup containing the culprit. This could explain why the filler choice rates are so much lower in the Illinois Pilot Program¹³ than in laboratory studies.

Investigator Bias***i) Conditions that Might Suppress or Enhance Investigator Bias***

Based on the observed difference in choice rates between the data from the Illinois Pilot Program¹⁴ and the predictions grounded in laboratory research, we wanted to examine whether specific variables that might make it easier or harder for investigators to influence witness choices (in the non-blind simultaneous lineup) had the predicted effects on choice rates. One such variable is the relationship that existed between the witness and the suspect prior to the crime. One might expect witnesses who knew the suspect prior to the crime would be more difficult for investigators to

influence than witnesses who were attempting to identify a stranger. Therefore, not only should strangers be identified less often than acquaintances (regardless of lineup procedure), but the investigators conducting simultaneous lineups should be able to influence witnesses to pick the suspect more often and the fillers less often when the suspect and witness were strangers. This should not occur with blind, sequential lineups.

We examined the prior relationship predictions in photo lineups, as most live lineups contained stranger relationships. With simultaneous photo lineups, 90.3 percent of the witnesses chose the suspect when a prior relationship existed but only 53.6 percent chose the suspect when they were strangers. With sequential photo lineups, these percentages were 76.3 percent

[A]cross all three jurisdictions, the known error rate (i.e. foil identifications) was higher with sequential than simultaneous lineups, and suspect choice rates were higher in simultaneous than sequential lineups for two out of three of the jurisdictions. It is clear that the results from this field program are in direct contradiction to the generalizations from laboratory research findings.

and 43.8 percent respectively. Thus, the difference in choice rates between sequential (blind and no influence) and simultaneous (with influence) lineups was larger for the acquaintance choices (76.3 percent v. 90.3 percent) than for the stranger choices (43.8 percent v. 53.6 percent), exactly opposite to prediction.

Considering filler choices from simultaneous photo lineups, none of the witnesses or victims chose a filler when a prior relationship existed; but, 1.3 percent of the witness/victims chose a filler when they were strangers. With sequential photo lineups, these

(Ebbesen/Finklea, continued on page 12)

SYMPOSIUM ISSUE

(Ebbesen/Finklea, *continued from page 11*)

percentages were 2.6 percent and 9.4 percent respectively. As expected, fillers were chosen more often when a “stranger” relationship existed in both lineup types, but were investigators who conducted simultaneous lineups better able to influence witnesses to avoid choosing fillers in stranger lineups (compared to acquaintance lineups) than were investigators who conducted sequential lineups? The shift from sequential to simultaneous lineups caused a small decrease (from 2.6 percent to 0) in filler choices when the suspect was an acquaintance but a bigger decrease (from 9.4 percent to 1.3 percent) when the suspect was a stranger. These results are consistent with the investigator bias explanation.

Consequently, although the filler choice rates for strangers and acquaintances might be explained by investigator bias, the suspect choice rates are inconsistent with this hypothesis. Since the investigator influence hypothesis assumes that investigators would be simultaneously directing witnesses away from fillers and towards suspects, the pattern of results seems inconsistent with the investigator bias explanation.

We also analyzed whether the status of witnesses (as a victim of the criminal act or simply a witness to the action) had any effect of choice rates. Because the consequences of making a choice are different for the two types of witnesses, we might expect those who were victims of the crime to be more likely to make a selection purely for the sake of conviction. Investigators could take advantage of this tendency when they know who the suspect is in the lineup, an argument for a potential benefit of instituting a (sequential) double-blind procedure. If so, we may expect victims to be less likely to select foils and more likely to choose suspects, but only when presented

with the simultaneous procedure—a procedure in which the investigators knew who the suspect was. Results of the analysis are inconsistent with this view. Given a simultaneous lineup procedure, victims and witnesses selected suspects at equal rates (63.49 percent and 67.69 percent, respectively) and chose fillers at equal rates (2.07 percent, 1.83 percent). The same pattern held in sequential lineups for victims and witnesses selecting suspects (51.17 percent, 53.96 percent) and fillers (8.72 percent, 6.47 percent). In essence, the effect of lineup procedure on choice rates was unchanged for victims and witnesses.

ii) Witness Confidence

The double-blind procedure was included in the Illinois Pilot Program,¹⁵ in part, because researchers have suggested that without it, there is a possibility that investigators may consciously or inadvertently influence witness selections from a lineup.¹⁶ Were this to happen, we might expect those witnesses who agreed with the investigator to be more confident that they were right in selecting the suspect (the same person the investigator believed was guilty) and less confident when they disagreed with the investigator and selected a filler. To examine this notion, we first analyzed suspect and filler choice rates as a function of the confidence that witnesses expressed in their identifications (Table 2). Confidence could be assessed for 31 percent of the simultaneous lineups and 63 percent of the sequential lineups based on the investigators’ written assessments of witness confidence at the time the lineup identification was conducted. High and moderate confidence choices were more associated with higher suspect choice rates than were low confidence choices. Moderate and low confidence choices were associated

(Ebbesen/Finklea, *continued on page 27*)

Table 2. Number and Percent of Suspect and Filler Choices as a Function of Expressed Confidence for all Lineups with Known Suspect Structure

Confidence	Number of Choices			Percent of Choices	
	Suspect	Filler	No Choice	Suspect	Filler
High	186	7	64	72.4	2.7
Moderate	32	9	2	74.4	20.9
Low	7	4	18	24.1	13.8
Not Known	204	12	174	52.3	3.1

SYMPOSIUM ISSUE**The Street, The Lab, The Courtroom,
The Meeting Room**

James M. Doyle¹, Steven Penrod, Ph.D.,
Margaret Bull Kovera, Ph.D. and Jennifer
Dysart, Ph.D.

The Mecklenburg Report² documenting the results of the Illinois Pilot Program on Sequential Double-Blind Identification Procedures will frustrate the most violent partisans on both sides of the debate over the future of eyewitness investigations.

Sadly, the Mecklenburg Report will also disappoint a broad audience of practitioners in the middle that hoped for guidance—for something to do (or avoid doing) to minimize the number of imprisoned innocents and untouched criminals that the DNA exoneration cases warn us eyewitness memory can produce. The Mecklenburg Report represents a taxing effort, its author and the participating officers and departments deserve our thanks, but the Report does not succeed in combining the perspectives of the street investigator, the laboratory scientist and the courtroom litigator into a working synthesis. In the Report's aftermath, it is clearer than ever that all three perspectives, their potentials and their limitations must be recognized before there can be a basis for action informed by science.

If the Mecklenburg Report convinces the criminal justice system's practitioners—investigators, prosecutors, defenders and judges—that they cannot wait around for legislatures to act, but must get themselves to the table together, engage the scientists, and work to find answers, then it can be a positive contribution. But until that happens, the Mecklenburg Report will leave us not far from where we were when the National Institute of Justice issued its path-breaking *Eyewitness Evidence: A Guide for Law Enforcement*³ (“the Guide”) in 1999. We still have a substantial body of laboratory results arguing for the procedures that the Guide identified as good (pre-lineup instructions), or good but not preferred (sequential lineup display), or simply potential (double-blind) options. We now have a number of satisfied jurisdictions around the country (including New Jersey, Boston and Minneapolis) that have instituted double-

blind, sequential routines with apparent success. Even so, the Guide did place the burden of proving the superiority of those procedures in operation on their advocates, and the Mecklenburg Report's numbers certainly do not lighten that burden. In fact, at least on the surface, the Report's numbers seem to tend in the opposite direction: against innovation. But if we look beneath the surface, we find that even if the numerical results noted in the Mecklenburg Report's field study had been *reversed* the Report *still* would not have proved the superiority of double-blind sequential procedures. The Mecklenburg Report reveals a study that simply was not set up to test under scientific control either double-blind, or sequential procedures. Nor did it test scientifically any differences between blind and not-blind simultaneous procedures.

These gaps are doubly unfortunate because five years from now we will not be handling eyewitness identifications in the same way that we handle them now. Our arrays of dog-eared mug shots and hastily improvised station house lineups are certain to be supplemented—and are almost certainly doomed—by a digital revolution that promises us quick, cheap, convenient and comprehensively documented identification procedures. We will have—some departments now have—photo lineup capability on laptops in squad cars. The capacity to present photo-arrays on Palm Pilots has already been studied in the labs. There soon will be many alternatives to dragging victims to the precinct house in the middle of the night and hiring line-up fillers from the homeless shelters in order to test witness memory.

But to say we will have new equipment doesn't tell us what we should do with it. What should we show the witnesses on our laptops? “Sequential” displays? How should we show it? With “double-blind” techniques? What is the best procedure for the future? The Mecklenburg Report does not really answer these questions; in fact, the study it recounts does not really ask them.

In hindsight it is clear that the Mecklenburg Report reveals a crippling misunderstanding at the heart of the field study it describes.

The Illinois Legislature issued a directive to pursue a specific goal: compare a traditional technique of eyewitness evidence gathering (the “simultaneous”

(Doyle, *continued on page 14*)

(Doyle, continued from page 13)

display of suspect and fillers by an officer who is aware of which lineup member is the suspect) with a proposed improvement (the “sequential” display of suspect and fillers by a “double-blind” administrator). But the Mecklenburg Report is pervaded by an unexplained determination to treat the Legislature’s statement of a goal as if it dictated a *method*. The failure to acknowledge the distinction between goal and method affected not only the Report on the study, but the design of the study itself, and it imposed serious handicaps. The muddle of method and goal explains why, in the words of United States Attorney Patrick Fitzgerald, the study “raises more questions than it answers.”⁴

In fairness to the Illinois Legislature, it did what it could to signal that the studies of eyewitness procedures that should be conducted were *not* the crude “traditional v. double-blind sequential” test the Mecklenburg Report describes. The Legislature sought an empirical answer to an empirical question by the use of study instruments: “[D]esigned to elicit information for comparative evaluation purposes, and . . . consistent with objective scientific research methodology.”⁵ An appropriate objective scientific research methodology exists, but the Mecklenburg Report shows plainly that the study it discusses stopped short of applying that methodology.

Four psychologists are mentioned prominently in the Mecklenburg Report. Two, (Dr. Nancy Steblay and Dr. Gary Wells) are bitterly critical of the report and allege that their participation is exaggerated by the Report’s author; two (Dr. Ebbe Ebbesen and Dr. Roy Malpass) consider the Mecklenburg Report on the field study to be a valuable document. But *all four* of these scientists disclaim any responsibility for *designing* the study. In fact, the study, like the Report, is the product of a single hard-working lawyer for the Chicago Police without formal training in social science methods. The differences between legal and social science practice show. For example, the Report describes as “random” assignment methods which in the legal world might be accepted as meriting the term, but which no social scientist would recognize as true random assignment. The result of this absence of a science-based design is that the Mecklenburg Report forfeits lessons that a truly scientific approach might

have taught. Because the requirements of scientific methodology were not imposed in the design of the study, we now know much less than we could.

To begin with, a comparison of the new “double-blind sequential” photo arrays with traditional “simultaneous, not-blind” lineups, which merely lays the two side-by-side, could never have been informative in any scientific way. To properly assess the “sequential” photo-arrays against “simultaneous” procedures either *both* “simultaneous” and “sequential” would have to be “not-blind,” or *both* would have to be double-blind—only then could we gauge which factor was creating the effect we see. To properly weigh the impact of “double-blind” procedures *both* simultaneous and sequential procedures would have to be run in “double-blind” and “not-blind” conditions

The Mecklenburg Report speaks for many when it suggests that we expand our inquiries and address further questions. But, real improvements in justice system processes based on science will only occur if cops, prosecutors and defenders take responsibility for framing the right questions informed from the beginning by scientific advice.

before the impact of “blindness” on investigations could be assessed. In other words, a fatal “confound” is built into the design of the Report, making it impossible as a matter of method, to retrieve authoritative answers to the question the Legislature posed.

Besides, treating the question posed by the Legislature as a methodological directive while ignoring the Legislature’s wish that “scientific methodologies” govern the study hopelessly entangled the operational issues of what is feasible on the street or in the precinct with the reliability research issue of whether the new procedures are worth doing in the first place.

For operational purposes it was natural for the Report to use “suspect hits” as a proxy for “correct identifications.” A radical decline in “suspect hits” in

(Doyle, continued on page 15)

SYMPOSIUM ISSUE*(Doyle, continued from page 14)*

double-blind sequential procedures would indicate at least in a rough and ready way a very awkward disconnect between that particular identification technique and police field operations. But it is important to confine the “suspect hit” criteria to its operational significance. Obviously, if we were satisfied with every procedure that yields a 100 percent rate of “suspect hits” we would never have undertaken the enterprise in the first place; we would have simply agreed to regard the dozens of DNA exonerations (every one of them based on a “suspect hit” which seemed “corroborated”) and the dozens of active criminals who escaped justice in those cases as an inevitable cost of doing business. “Suspect hits” can tell us important things about operations, but—even in the Hennepin County field study, where the results were radically different from those recited in two of the three Mecklenburg Report jurisdictions—they tell us very little about the reliability questions at the heart of the issue of procedural superiority.

Lab methods have their own limitations, and there is a danger of unintentionally imposing those limitations if we undertake “lab-like” studies in the field.

The specific limitation that concerns us here is not the worry that in the real world crime situation human memory operates in a qualitatively different way. There is no evidence for that fear. In studies pre-dating the Report, the rate of “filler ID’s” in the lab and in the real world seemed to match up fairly closely. In the occasional hyper-realistic laboratory study, such as Dr. Charles Morgan’s controlled study of special forces troops who were asked to identify their interrogators after a high stress interrogation (more than half identified a “filler” in conventional simultaneous arrays) the results, again, are consistent with both the more conventional lab setting experiments and with the scattered field results from the United States and the United Kingdom.⁶ But even while we acknowledge that the lab studies and the field studies are examining the same processes of human memory, we have to remember that they do so in different contexts.

The Mecklenburg Report’s most intriguing results are its account of a “zero” rate of filler identifications in two of three jurisdictions. These results are unique among existing studies in the lab and in the field. How

did this happen? What does it mean? Unfortunately, the design of the study and the Mecklenburg Report’s recounting of it leave open a quite simple and obvious explanation: the failure to account for a fundamental difference between lab life and street life.

When the lab scientists study the efficacy of an identification technique, the single “simultaneous” or “sequential” test they scrutinize is almost always the witness’s *first* attempt at the identification of a stranger-perpetrator. This places the focus on the most influential (and therefore dangerous) encounter, but it does not automatically duplicate typical real-life practice; in real life a witness’s live line-up performance is only one episode in the witness’s career in the criminal justice system. In real life—and in the experience of the witnesses depicted in the Mecklenburg Report—a “live” lineup experience can be (in most places, usually is) preceded by a show-up, a “drive-by”, or by a photo-array. To treat the rate of suspect identifications attained in *first* attempts in laboratories and *third* attempts in the field (by witnesses who were, in effect, pre-tested on a show-up and a photo-array) as equivalent doesn’t just compare apples to oranges; it compares apples to automobiles. It isn’t particularly surprising if third attempts by pre-tested witnesses (i.e., after two successful attempts and the dismissal of all of the unsuccessful witnesses) to identify a suspect in the field lead to fewer “filler ID’s.”

We don’t know from the Mecklenburg Report that this happened, but unfortunately we can’t know that it *didn’t* happen, because the witnesses’ history in the investigation is not recorded or reported. The Mecklenburg Report treats the field results as if they were the lab results, but the study under examination did not follow the scientific tradition of recording experimental data, and so it failed to capture data that the lab would have noted as a matter of routine experimental design. How many of these eyewitnesses identified fillers in initial field procedures? How many of these were filtered out of the process before the subsequent, reported lineups? How many witnesses in the subsequent lineups were performing a *confirming* recognition task following a successful suspect identification in a show-up or an array? Either of these features is at least as likely to have affected the suspect/filler identification rates as might wholesale

(Doyle, continued on page 16)

SYMPOSIUM ISSUE

(Doyle, continued from page 15)

police “tipping” of witnesses by police. (Although the Mecklenburg Report persistently invokes the straw man of sinister allegations of police misconduct, in fact no one claims either that intentional police misconduct is the problem in the DNA eyewitness exoneration cases, or that if it *were* the problem, procedural reforms would be a silver-bullet solution to all intentional “framing” of suspects.) If people are cheating, they will continue to cheat, whatever procedures are adopted. But, as things now stand we are provided with no authoritative refutation of dark speculations about of police “tipping” because a wide range of data points—for example, the number of “low confidence” filler identifications and the number of failures to identify—were not captured in the study’s design.

Operations and reliability are muddled in a different way when the “double-blind” technique is at issue. If “double-blind” procedures add something to accuracy, then implementing double-blind procedures—as jurisdictions in Wisconsin, Massachusetts, New Jersey and elsewhere have done—becomes a question of police ingenuity, commitment and leadership in surmounting operational challenges. But, if we ask the police their opinion of the “double-blind” approach before they are persuaded that it can contribute to accuracy, the police can’t be blamed for accepting inconvenience and unfamiliarity as sufficient answer. Besides, when the police are not invited to participate in the design of the specific local double-blind sequential technique but are simply presented with a “take it or leave it” version in informal oral training an opportunity

The Mecklenburg Report represents a taxing effort, its author and the participating officers and departments deserve our thanks, but the Report does not succeed in combining the perspectives of the street investigator, the laboratory scientist and the courtroom litigator into a working synthesis.

to confuse the performance is created and an opportunity to exploit police expertise is lost.

The Mecklenburg Report speaks for many when it suggests that we expand our inquiries and address further questions. But, real improvement in justice system processes based on science will only occur if the cops, prosecutors and defenders take responsibility for framing the right questions informed from the beginning by scientific advice. The system’s practitioners not only have to take responsibility for integrating science into practice; they have to take responsibility for doing it *together*.

This will cause some discomfort. The Mecklenburg Report documents a field study that followed the more normal course of reform efforts within the system: One actor or another is charged with (or pro-actively assumes) responsibility for mobilizing one scientific advance or another, chooses its own scientists, closely holds the information developed and makes (or foregoes) reforms. This is not the only way.

Behind their adversarial routines, all criminal justice practitioners share a common enemy—the innocent defendant. No one wants the innocent in the system. The police do not want to waste their time on the innocent while the guilty go free to prey on new victims; the prosecutors realize that highly publicized exonerations in the cases they should have lost will later cost them the cases they should win. Maybe young defense lawyers go to law school with dreams of defending the innocent, but experienced defense lawyers see defending an innocent—particularly in an eyewitness case—as a nightmare. Double-blind, sequential lineup procedures—if they work to keep the innocent out of the system—are to everyone’s advantage, and they should get a genuine scientific test for that reason alone.

It is also worth remembering that the question of eyewitness identification reform is not an all-or-nothing matter. Sophisticated police departments might, after testing, decide that some crime situations (for example, where there is a substantial amount of corroborating information) call for traditional methods of identification, while other, shakier, cases call for the more cautious, conservative double-blind sequential approach. “Double-blind” administration on its own (even if “simultaneous”) also serves important law

(Doyle, continued on page 31)

SYMPOSIUM ISSUE

Getting to Truth Before It Falls into the Hands of the Lawyers: Pursuing Accuracy in Criminal Cases

James B. Zagel¹

Societies have always wrestled with the overall question of the reliability of witnesses and, even now, when the legal rules are mostly settled, we still worry about perjury, mistakes, delusions and the integrity of memory. At issue today is the relatively small subset of the witness problem—just that one moment when the witness points to one person and says that is the person whose conduct I have described. For most witnesses to crime, the phrase “That’s the man” is shortest part of the story they offer. Concern about its accuracy has been with us for centuries.

If solutions were easily found, this would not be an age old concern. Be wary of those who, with great confidence, offer the miracle cure to a problem we all recognize. The results of the Illinois double-blind eyewitness pilot program offer a vivid example of why what some think obvious is often not so. There is another point here; failure teaches as much or more than does success and we ought not to turn our back on any enterprise that seeks to make our investigations and adjudications better. The great value of tests, like the one we discuss herein, is they keep us from a terrible kind of optimism that, once disappointed, can lead us to abandoning the search for something better.

What Is It We Are Trying To Repair and Why?

Our world of arrest, prosecution and defense has changed. The idea that truth arises out of trial in an adversary system is still with us but mostly in theory not practice. Plea bargains are the dominant mode. My colleague, Judge Lynch in New York, has accurately described the process this way:

“[T]he prosecutor ... is the central adjudicator of facts ... arbiter of ... legal issues and of the appropriate sentence to be imposed. Potential defenses are presented by the defendant ... to a prosecutor, who assesses their factual accuracy and then decides the charge of which the defendant should be adjudged guilty...”²

This is a far cry from what the Supreme Court

envisioned when it began to emphasize 75 years ago that defense counsel at trial was essential to getting at the truth.³ And the place, they thought, where truth was to be found was trial where it was judge and jury, after hearing prosecution and defense, not prosecutor, who decided the outcome. This reliance on defense counsel to help us get to the truth was a key element in the first cases in which the Supreme Court sought to bring constitutional regulation to eyewitness identification.⁴

But defenders are not duty bound to see that the truth comes out. If the client is guilty they are obliged



It is critical to establish an identification before the eyewitness is subject to outside suggestion and a false identification is made. A false identification often has irreversible consequences.

to use all legally permissible means to see that the truth does not come out. This became particularly clear when we thought about what a lawyer should do at a lineup. Suppose the client tells his lawyer “Yeah, I stole the stuff but I’m sure no one saw me inside.” Then the lawyer sees his client in a proposed line-up of seven, six of whom are Hispanic, and his client is the only blond white man in the group. Does

(Zagel, continued on page 18)

SYMPOSIUM ISSUE

(Zagel, continued from page 17)

counsel want a fairer or more accurate lineup, one more likely to elicit the truth? If he gets one and the identification is made, he has deprived his client of a good argument at trial. What if the police ask the lawyer for suggestions to improve the lineup and agree that they will follow these suggestions? Does the lawyer improve the process to the detriment of his client? The dilemma here is stark because, unlike interrogation, the lawyer can not simply advise his client not to participate in the lineup. The identification procedure is going to happen. The lawyer is not authorized to decide simply that it is right to have a fair identification parade; the lawyer is only authorized to seek the kind of parade that is good for the client.

Many eventually accepted this state of affairs where, in a trial, getting to the truth was not the single overriding value.⁵ I think they did so for two reasons. First, the thought was that, in nearly all cases, the truth came out anyway. Second, there were important social values found in procedural fairness and in giving the defendant a meaningful role in his or her defense. The price of an occasional criminal going free was thought to be worth paying to achieve these good things. This tradeoff has always been controversial. It might not survive a public referendum.

The tradeoff also rests on premises that professionals find hard to accept. It is not easy to find scholars (though not so hard to find judges) who actually believed that trial was really a good way to get at the truth in hard cases; the scholarly defense of the system was based upon its service to other democratic values. And even where the adversarial system could work, it was dependent on having a skilled, adequately funded advocate on both sides of the case. This last condition was often unmet.

In the decades that followed the criminal procedure upheavals of the 1960s, there was a lukewarm to cool acceptance of the way criminal cases were handled: lukewarm to cool because we were in the midst of a rise in crime that lasted for decades and, only relatively recently, subsided; accepted because there was much in popular media,

shows like *NYPD Blue* and *Law and Order*, that portrayed a system that got the right result. The right result is the common result, but it is not because of our system of trials. It is because, in most cases, there is no serious question of guilt. The evidence is usually more than good enough and, if it isn't, the prosecutors frequently won't take a chance on the case. Perfection is unattainable, but getting it right in the largest percentage of cases is not good enough even if that percentage is in the high nineties. The consequences of error are too grave.

But still we toddled along with what we had. The volume of criminal cases, which was associated with the rise in crime and the relatively poorly funded defense services, led to fewer trials and what Judge Lynch called an administrative system of criminal justice. We would be that way today but for recent events.

The public, as opposed to the defenders of the world, was generally worried only about the guilty going free. New science and today's news made them worry, at least a little, about the innocent being found guilty.

For this reason the importance of finding the truth about guilt is valued more highly today than it has been in many years. We might be in the midst of a tectonic shift in perspective about crime, investigation and the accused.

Some of this might seem strange to say to the public. Haven't we always thought that the end of the criminal justice system was to find out the truth? Most people did, but no one who labors in this field believes that truth is always revealed or acted upon. There are unjustified convictions and unjustified acquittals. Even under the better practices that we will someday have, we will never reach perfection because the truth is elusive, often beyond the ability of humans to discover. But we won't stop prosecuting. Crime has a devastating effect on its victims and a large effect on the society in which we live. We have never decided to leave the guilty or the innocent to the judgment of heaven. We ought then to do the best we can to lock up all the guilty and free all the innocent, knowing that we will sometimes fail.

(Zagel, continued on page 19)

SYMPOSIUM ISSUE

(Zagel, continued from page 18)

What Has Led Us to Try Now?

The short answer to this question is that there is no sudden change, just a gaudy tipping point. The shift to concern for truth has been building for a while.

Start with *Miranda v. Arizona*⁶ and the de-emphasis on admissions of guilt. *Miranda* can now be read as an implied endorsement of the reliability of eyewitnesses and, perhaps, of the very forensic evidence that is now under attack, say, bite marks, and even of reliable evidence excludible under the Fourth Amendment. The Court understood it was holding that even voluntary confessions, whose truth value was unquestioned, were to be excluded from evidence. This de-emphasis of truth and re-emphasis on procedural protections did not last very long. The Supreme Court limited the scope of *Miranda*.⁷ The Court (in an opinion written by Justice Thurgood Marshall) also decided to permit the police to use deception to induce confessions⁸ and narrowed the scope of the exclusionary rules based on the Fourth Amendment.⁹



The idea of science in law enforcement is relatively new. Identification by fingerprinting, for example, was not broadly accepted in the United States until the 1920s.

Despite this renewed endorsement of the value of confessions, there is no doubt that the Supreme Court, and every sane person for that matter, would prefer that guilt be determined by incontestable evidence like that found three times every week in New York, Miami and Las Vegas by an infallible corps of Crime Scene Investigators. It is science that created that gaudy tipping point.

The entry of science into the courtroom started at a very slow pace. Most of the earliest expert testimony from doctors and alienists was admitted into

evidence because the law permitted *anybody* to offer opinion evidence.

Fingerprints are a good example of the early evidence of experts. Fingerprints were used for identification in India in the 1850s.¹⁰ Written work appeared in 1881 and Galton's book was published in England in 1892.¹¹ By 1910, fingerprinting itself was in fairly wide police use.¹² Despite this, it was not until 1911 that a reviewing court approved its use, but broad acceptance did not come until the 1920s, and it was not until the 1940s that courts said the prosecution would no longer have to prove that no two fingerprints are alike.¹³ The course of admission of other forms of identification evidence was similar. It took time to get the courts to approve comparative micrography, microanalysis, questioned documents. Most of the first scientific evidence dealt with traces and marks which the jurors themselves could perceive—friction ridges, striations on bullets and so forth. As the twentieth century went on, and science itself began to deal with things not directly observable, the law began to take in serology, general chemistry and neutron activation analysis. In all these cases, though, one reason the courts moved slowly was the resistance of defense counsel to the admission of such evidence because it rarely served any purpose other than to incriminate their clients.

DNA evidence was accepted with amazing speed precisely because it could exonerate as well as incriminate. There were very few to fight tooth and nail against its admissibility because the prosecutor or defender who objected vigorously to DNA evidence knew that, in the next case, they might be offering that same evidence. DNA, too, came to the courts at a time when standards of what constituted reliable and valid science had become clearer. DNA analysis had the advantage of service as a tool in many sciences, not merely criminalistics. The broader use of DNA analysis meant the discipline had been critically reviewed by many more scientists than, say, fingerprints.

It is true that the advent of closer judicial scrutiny of expert witnesses in recent years¹⁴ has called into question much of the science that is offered in the courtroom, but the outcome of disputes about questioned document examination, serology, fiber analysis,

(Zagel, continued on page 20)