

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Court Review: The Journal of the American
Judges Association

American Judges Association

2012

Ready for the Psychologists: Learning from Eyewitness Errors

James M. Doyle

National Institute of Justice

Follow this and additional works at: <https://digitalcommons.unl.edu/ajacourtreview>

Doyle, James M., "Ready for the Psychologists: Learning from Eyewitness Errors" (2012). *Court Review: The Journal of the American Judges Association*. 379.
<https://digitalcommons.unl.edu/ajacourtreview/379>

This Article is brought to you for free and open access by the American Judges Association at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Court Review: The Journal of the American Judges Association by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Ready for the Psychologists: Learning from Eyewitness Errors

James M. Doyle

Over a century ago, Dean John Henry Wigmore published a famous demolition of pioneering psychologist Hugo Munsterberg in the *Illinois Law Review*.¹ Munsterberg had complained in his best seller, *On the Witness Stand*, that while other disciplines and professions were hustling to learn the lessons about eyewitness memory that his new field of experimental psychology was beginning to teach, “the lawyer alone is obdurate.”² Munsterberg charged that the lawyers chose traditional primitive ignorance over scientific enlightenment. Wigmore could not sit still for that. His satirical response is still remembered by psychologists as the blood-thirsty slaughter of psychology as a discipline by the greatest evidence scholar that the Anglo-American tradition ever produced: a grisly paradigm of the kind of welcome social scientists should expect from the legal system and its practitioners. If this is what you get from the great Wigmore, researchers reasoned, just imagine the treatment you will receive from an ordinary legal tribesman.

Wigmore’s withering cross-examination of the wretched “Professor Muensterberg” in this article is so lengthy and so humiliating that there are moments when a slightly creepy sadistic pleasure seems to be animating the dean. But sadism wasn’t the problem. The problem was Wigmore’s cloddish professorial attempts at humor—Wigmore’s sarcasm created a misimpression that he tried to correct for the rest of his life. Wigmore did want to issue a call to order: to correct Munsterberg’s overstatements and to address Munsterberg’s misapprehensions about legal practice. But Wigmore was far from an enemy of psychology as a discipline; he was actually one of psychology’s earliest advocates, the best legal friend that psychology had.

The real purpose of Wigmore’s article was to illuminate the potential in a law and psychology relationship and to throw his prestige behind its inception. Wigmore’s goal was to herald the day when the lawyers and psychologists could move forward in “a friendly and energetic alliance in the noble cause of justice.”³ Yes, Munsterberg had jumped the gun in announcing the immediate utility of such an alliance; Wigmore thought that was still on the distant horizon. Even so, Wigmore looked forward to the day of its arrival, and he was confident that day would come. “When the psychologists are ready for the courts,” he announced in a subsequent piece, “the courts will

be ready for the psychologists.”⁴

Various signs and portents—among them, this special issue—indicate that the courts finally are ready to mobilize the lessons taught by Munsterberg and his heirs; or at least that the courts are ready to take steps to get ready. This is an important moment in the vexed history of the law and eyewitness psychology relationship.

To understand where we are it helps to understand a little about both how we got here and where we could be going. The “friendly and energetic alliance” will have more than one path to choose from as it moves ahead. The path that realizes the fullest potential of the alliance is not the most obvious path, and finding it will require a new examination of the deep nature of the catalyst—the devastating catalog of DNA exonerations in eyewitness cases—that is pushing us forward.

Something more than a minor adjustment to judicial practice is called for here: this is an opportunity for judicial leadership.

DIAGNOSES AND PROBABILITIES

Hugo Munsterberg’s pronouncements on the usefulness of contemporary psychology were overconfident and premature, but they were also remarkably prescient in anticipating future research. Munsterberg began to explain some of eyewitness memory’s mechanisms and some of its particular dangers. He showed, for example, that humans do not have a permanent stable memory capacity like a videotape or a DVR available to be summoned for accurate replay whenever required. He showed that memory was malleable and reconstructive.⁵ He also showed how forensic evaluations of memory evidence could go astray. For example, he showed that a witness’s confidence was an unreliable indicator of the witness’s accuracy.⁶ But for current purposes, Munsterberg’s method was as important as his findings.

Munsterberg’s signature tool (at least for public display) was the staged demonstration. A man interrupts a lecture; he yells; he fires a gun; later, the audience of eyewitnesses is questioned about the event. Next, inaccuracies in the audience’s responses are totaled. The number of errors in the witnesses’ responses is shocking. Lots of eyewitnesses make lots of mistakes. These results grabbed attention, but they were not terribly useful for the legal system. They indicated that there were many mistakes,

Footnotes

1. John H. Wigmore, *Professor Muensterberg and the Psychology of Testimony: Being a Report of the Case of Cokestone v. Muensterberg*, 3 ILL. L. REV. 399 (1909).
2. HUGO MUNSTERBERG, *ON THE WITNESS STAND* (1908). The story of the encounter and its aftermath is told in some detail in JAMES M. DOYLE, *TRUE WITNESS: COPS, COURTS, SCIENCE, AND THE BATTLE AGAINST MISIDENTIFICATION* (2005).

3. Wigmore, *supra* note 1, at 406. In fact, Wigmore may have had a better grasp of contemporary psychological research than Munsterberg himself. See Brian H. Bornstein & Steven D. Penrod, *Hugo Who? G.F. Arnold’s Alternative Early Approach to Psychology and Law*, 22 APPLIED COGNITIVE PSYCHOL. 759, 762 (2008).
4. JOHN HENRY WIGMORE, *THE SCIENCE OF JUDICIAL PROOF* (1937).
5. MUNSTERBERG, *supra* note 2, at 60.
6. *Id.*, at 56.

and they argued for an increased general skepticism about eyewitness accounts. But, as Wigmore pointed out, the legal system's concern is not with the general reliability of witnesses as a class; it is with the reliability of particular verdicts in individual cases. The legal problem arose in separating the mistaken from the correct—not the rate of mistakes, but their distribution. There, Munsterberg had little or nothing practical to offer.

When Robert Buckhout picked up Munsterberg's fallen banner in the 1970s, he relied on a modernized version of the same approach.⁷ For example, he induced a New York television station to broadcast a staged crime and invite viewers to make choices from a staged lineup. The number of correct identifications this process yielded was lower than would have been achieved by random guessing.⁸ But while his method may have been similar, Buckhout's temperament was very different from Munsterberg's. Munsterberg was an academic who retreated when faced with Wigmore's onslaught. Buckhout knew his science, but he was a happy warrior, a cheerful agitator who carried the battle into the courts and into the popular media. He not only accepted opposition, he gloried in it.

He published an accessible survey article on eyewitness error in *Scientific American*.⁹ He testified on the unreliability of eyewitness testimony in the trial of California radical Angela Davis and was instrumental in winning her acquittal. He seized every opportunity to comment in the media (for example, opining on the case of a butcher identifying his own pork chops from a pork-chop lineup) where the lessons of eyewitness psychology could be taught. His science was aligned with his politics. He believed that criminal defendants, particularly poor and minority indigent defendants, were getting screwed by the legal system's complacent reliance on an antique view of how memory worked. He made an enormous impact, and he almost immediately rallied two groups of partners.

The first group was a cohort of idealistic younger psychologists, like Elizabeth Loftus, who were anxious to see their science have an impact in the world. Loftus attacked the eyewitness issue in a radically different way: she "did science" in the form of rigorously controlled experiments, changing one variable while holding all others constant. The results she began to produce were striking. She showed, for example, that when questions about a white barn were introduced into interrogations of witnesses who had viewed a film of an auto accident, over 20% of those viewers later reported seeing a white barn although in fact there had been no white barn in the film.¹⁰ This was a crucial finding for eyewitness cases: it showed that eyewitness memory not only decayed, but also *changed*. It showed how a witness could not only forget the right man but also—after being unknowingly influenced by viewing mug shots or show-ups (which operate as "post-event information" like the white barn in an interview question)—could remember the *wrong* man.

Loftus's findings mounted quickly, and they went to the heart of the eyewitness experience. Taken together they indicated that in an eyewitness case, the memory of the witness is for all practical purposes the scene of the crime. They showed that memory evidence was in effect "trace evidence": difficult to collect, easy to contaminate, but impossible to test for contamination after any contamination has occurred. At the same time, Loftus's scrupulous scientific methods were winning her work admission to the blue-ribbon, peer-reviewed academic journals, and encouraging younger academic psychologists to extend and challenge her research. You could study eyewitnesses and have a scholarly career. Experimental findings such as Loftus's (unlike the demonstrations of Munsterberg and Buckhout) could be replicated or falsified. The number of published studies multiplied.¹¹

And at this point, Buckhout's second group of recruits, the desperate criminal defense lawyers, joined in. Buckhout's testimony in the Angela Davis case got their attention, and his *Scientific American* article quickly circulated through the defense bar. Elizabeth Loftus published her popular general audience account of eyewitness science, *Eyewitness Testimony*,¹² at about this time, and that was buttressed by an influential *Stanford Law Review* comment written by Frederick Woocher (a trained psychologist, then in law school), which provided a blueprint for arguments for conveying psychological science through expert witnesses. Defense lawyers began to demand the admission of expert testimony by Loftus, Buckhout, and their colleagues, aimed at debunking faith in eyewitness evidence.

This point of entry was bad luck for anyone who hoped for a "friendly and energetic alliance." That wasn't obvious at the time. Persistent litigation over admissibility did help to keep the issue of eyewitness science alive in the courts, and feedback from skeptical courts did help to provoke new, better-targeted research. But these benefits came at a steep price.

The initial environment has affected discussions of eyewitness science ever since. Admissibility questions arise at the most acutely adversarial moments of the criminal process, and their resolution (at least in the eyes of the advocates) may determine who wins and who loses. Prosecutors—goaded by inflammatory rhetoric from Buckhout—quickly denounced eyewitness findings as enemy pseudoscience: a trick designed to let criminals go free by unnerving credulous lay jurors and sliming all eyewitnesses, most of who were right, and many of who were crime victims. For many prosecutors—then and

Loftus's research "showed how a witness could not only forget the right man but also... could remember the wrong man."

7. DOYLE, *supra* note 2, at 49-68, discusses Buckhout's history and influence.

8. Robert Buckhout, *Nearly 2,000 Witnesses Can Be Wrong*, 2 SOCIAL ACTION & L. 7 (1975).

9. Robert Buckhout, *Eyewitness Testimony*, 231 SCI. AM. 23 (1974).

10. Elizabeth F. Loftus & John C. Palmer, *Reconstruction of Automobile*

Destruction: An Example of the Interaction Between Language and Memory, 13 J. VERBAL LEARNING & VERBAL BEHAV. 585 (1974).

11. BRIAN CUTLER & STEVEN PENROD, *MISTAKEN IDENTIFICATION: THE EYEWITNESS, PSYCHOLOGY AND THE LAW* (1995).

12. ELIZABETH LOFTUS, *EYEWITNESS TESTIMONY* (1979).

“Wells argued that preventing mistakes [during the] investigation would be better than trying to catch mistakes....”

now—eyewitness science is simply a shield for the guilty. For many judges, the cumulative price of the skirmishing over marginally interesting science the experts offered seemed enormous in terms of hours, dollars, and distended docket backlog.

While the battles over admissibility of expert testimony continued to grind on, another of Buckhout’s recruits, Gary Wells,

was engineering a paradigm shift.¹³ Wells admired Loftus and accepted her findings as good science, but he also pointed out their limited utility.

Precisely because Loftus was a scrupulous scientist, she isolated and studied a single factor (e.g., the wording of a question, the stress of the event, the presence of post-event information) at a time. Wells noted that these studies yielded statistical results that could tell you what happened eight times out of ten, but could not tell you whether *this* case was one of the eight, or one of the two. Even worse, every criminal event incorporates many factors, not just one, and there was no science-based mechanism for combining these factors and assessing their interactions. From Wells’s point of view, offering post-hoc diagnosis of eyewitness error from the witness stand was the wrong way to mobilize the solid (but inherently probability-based) science that Loftus and a generation of their colleagues were producing.¹⁴

Wells successfully argued for the new orientation that has dominated criminal justice policy discussions about eyewitnesses for the past decade. He noted that some factors Loftus had studied (e.g., lighting, age of witness, stress of event) are not under the criminal justice system’s control. He called these “estimator variables.” But he also noted that there were other factors (e.g., lineup construction, lineup administration, witness interview technique) that the system’s actors *do* have power over. If you understood how these “system variables” could be modernized, you could reduce the rate of error. Wells argued that preventing mistakes by identifying new best practices in investigation would be better than trying to catch mistakes from the witness stand after they happened. A torrent of research followed, exploring and refining new elements of “system-variable” design. The task of psychological science in this conception was the prevention of eyewitness errors as evidence was being produced, not the retrospective inspection of eyewitness testimony to see if an error had occurred. That research has now coalesced around the “double-blind sequential” photo-

array and lineup protocol discussed later in this issue.

Then, just as that research matured, the DNA exoneration cases arrived. The eyewitness cases dominated the lists of wrongful convictions; the system-variable research was well developed, and its salience was immediately obvious. Influential actors such as Attorney General Janet Reno were eager to apply the researchers’ lessons.¹⁵ Expert-witness litigation does continue, and a gradual but definite trend toward the admission of eyewitness expert testimony in trials has gained momentum in the courts.¹⁶ But the policy conversation has turned toward prevention: toward the design of system-variable “best practice” reforms of lineup and other investigative procedures. An accelerating wave of jurisdictions has been adopting the science-based eyewitness-evidence protocols.

If this is where we are, then where are we going? The answer to that question will depend in part on how we understand the lessons of the DNA eyewitness exoneration cases.

THE WRONG MAN AND THE WRONG PATIENT

Wigmore’s “friendly and energetic alliance” received a dramatic push forward from the exoneration cases, but it would be a mistake to settle for the most obvious lessons that the eyewitness wrongful convictions seem to offer.

Smalarz and Wells are not wrong when they write that “[a]n increasingly strong case can be made for the argument that mistaken-eyewitness identification is the primary cause of the conviction of the innocent in the United States,”¹⁷ but their familiar formulation uses “cause” in a shorthand sense that may mask both the complexity of the issue and the opportunities for mobilizing science in reform that the collision of eyewitness psychology and the DNA exonerations provide.

One very good way to see those complexities and opportunities is to examine contemporary medicine’s encounter with its own version of the problem.

Just as the criminal justice system is haunted by the fact that it sometimes convicts the wrong man, medicine is haunted by the fact that it sometimes operates on the wrong patient. But when modern medical researchers began to look carefully into wrong-patient events, they uncovered surprising insights. For example, one intensive examination of a wrong-patient surgery discovered not just one but at least seventeen errors. The patient’s face was draped so that the physicians could not see it; a resident left the lab assuming the attending physician had ordered the invasive surgery without telling him; conflicting charts were overlooked; and contradictory patient stickers were ignored. But the crucial point for the researchers was that *no single one of the seventeen errors they catalogued could have caused the wrong-patient surgery by itself.*¹⁸

13. DOYLE, *supra* note 2, at 142-167.

14. Gary L. Wells, *Applied Eyewitness-Testimony Research: System Variables and Estimator Variables*, 36 J. PERSONALITY & SOC. PSYCHOL. 1546 (1978).

15. DOYLE, *supra* note 2, at 163-170.

16. See generally ELIZABETH LOFTUS, JAMES M. DOYLE & JENNIFER DYSART, *EYEWITNESS TESTIMONY: CIVIL AND CRIMINAL* (4th ed. 2007) at 369-377.

17. Laura Smalarz & Gary L. Wells, *Eyewitness Identification Evidence:*

Scientific Advances and the New Burden on Trial Judges, 48 CT. REV. 14, 14 (2012) (this issue). Recent investigations of non-DNA exonerations indicate that false testimony cases may outnumber eyewitness cases in certain varieties of false convictions. NATIONAL REGISTRY OF EXONERATIONS, <http://www.law.umich.edu/special/exoneration/Pages/about.aspx> (last visited July 9, 2012).

18. Mark R. Chassin & Elise C. Becher, *The Wrong Patient*, 136 ANNALS INTERNAL MED. 826, 829-31 (2002).

Analysis showed not only mistakes by individual doctors and nurses, but also latent systemic problems. Communications among staff were terrible; computer systems did not share information. When teams failed to function, no one was surprised or bothered because of a culture of low expectations that “led [staff] to conclude that these red flags signified not unusual, worrisome harbingers but rather mundane repetitions of the poor communication to which they become inured.”¹⁹ Deviations from good practice had become normal, and a tragedy resulted.

What this meant to medical reformers was that the lessons of closely studied events such as the Chernobyl meltdown and the space shuttle *Challenger* launch disaster could be applied to healthcare. Like those tragedies, the wrong-patient surgery was an “organizational accident.” No single error is sufficient to cause an organizational accident; the errors of many individuals (“active errors”) converge and interact with system weaknesses (“latent conditions”), increasing the likelihood that individual errors will do harm. The practitioners and organizations involved in these tragedies did not choose to make errors—they *drifted* into them.²⁰ The disasters required no villains; they involved normal people, doing normal work, in normal organizations.²¹ They suffered, in Charles Perrow’s memorable phrase, “normal accidents.”²² Like the *Challenger* launch decision, the medical tragedies were caused by “mistake[s] embedded in the banality of organizational life.”²³

These insights apply to a wrong-man conviction.²⁴ Our traditional wrongful-conviction narrative (the witness picked the wrong guy; the cops and the D.A. believed her; so did the jury) is not adequate. Nor is it adequate to isolate the performance of one operator or the imperfections one investigative technique employed in the case—for example, the traditional non-blind, simultaneous lineup—as either a sole cause or a silver-bullet solution.

Lots of things have to go wrong before the wrong man is convicted. Yes, the witness has to choose the wrong man from an array, but the police have to put him into the array in the first place and design the format of the array and the execution of the identification. Forensic evidence on the crime scene could have been overlooked or, although properly collected and tested in the lab, distorted in the courtroom presentation. Cell-phone records, Metrocard data, or other alibi information could have been ignored. Tunnel vision, augmented by clearance rate and caseload pressures from above, may have overwhelmed the investigators and the prosecutors. Poorly funded or untrained defense counsel may have failed to investigate alternative explanations or to execute effective cross-examination. The witness erred; the cops erred; the technicians erred; the prosecutors erred; the defense erred; the judge and the jury erred; the appellate court erred, too. No single one of these errors would have been enough without the others. The errors

combined and cascaded; *then* there was a tragedy—and a “no-villains” tragedy at that.

When we ask who is responsible for a wrongful conviction, the right answer is usually “everyone involved,” to one degree or another—if not by making a mistake, then by failing to catch one. And “everyone” includes not only cops and lawyers at the sharp end of the system, but also legislators, policymakers, funders, and appellate judges far from the scene of the events who dictated the conditions under which the sharp-end operators work. Look twice at the DNA-exposed wrongful convictions and you see that, as Charles Perrow noted, “[T]ime and again, the operator is confronted by unexpected and usually mysterious interactions among failures, [so that] saying that he should have ziggged instead of zagged is possible only after the fact.”²⁵ This is as true of a whole spectrum of criminal justice errors—mistaken releases, prisoners lost in prisons, and cold cases that stayed cold too long—as it is of wrongful convictions.

The habit of treating horrific wrongful convictions as single-cause events, and then totaling up, ranking, and prioritizing these causes, has produced useful innovations such as the double-blind sequential protocol and, in some places, has led those reforms to be integrated into practice, but it does not really engage the deeper nature of the problem. The solutions it has generated stop short of fundamentally improving future system reliability.

All new sets of best practices or checklists have to operationalized and executed, and they have to be maintained, monitored, evaluated, and perhaps junked and replaced when environments change or science advances. No new set of best identification practices can cover every circumstance, so an irreducible zone of discretion always survives, and operators are forced to manage life within that zone. From the moment it is written, every new checklist is under immediate and constant assault from clearance-rate pressure, docket-list backlog, and other environmental factors. “Drift” toward failure remains a threat to our new best practices just as it was to their discredited predecessors. No one had more checklists than NASA; NASA launched *Challenger* anyway.

Many tragic mishaps could never have been predicted (and cannot now be explained) by reference to the features of individual component parts. These tragedies are “emergent” events, results of the “greater than the sum of its parts” properties inherent in all systems.²⁶ Going “down and in” to find a broken component will not be enough to explain these happenings; we also have to go “up and out” to assess the envi-

“When we ask who is responsible for a wrongful conviction, the right answer is usually ‘everyone involved’...”

19. *Id.*, at 829-830.

20. SIDNEY DEKKER, *DRIFT INTO FAILURE* (2011).

21. SIDNEY DEKKER, *JUST CULTURE* (2007).

22. CHARLES PERROW, *NORMAL ACCIDENTS* (1984).

23. DIANE VAUGHAN, *THE CHALLENGER LAUNCH DECISION: RISKY TECHNOLOGY, CULTURE, AND DEVIANCE AT NASA*, xiv (1996).

24. This argument is made at greater length and in greater detail in James M. Doyle, *Learning from Error in American Criminal Justice*, 100 J. CRIM. L. & CRIMINOLOGY 109 (2010).

25. PERROW, *supra* note 22, at 9.

26. DEKKER, *supra* note 20.

“A modern approach to ‘best practices’ in collecting eyewitness-memory evidence is plainly called for....”

ronment that chose the component, allowed the component to fail, and made the failure catastrophic.²⁷ Making good design choices between alternative single components of the criminal process (e.g., between “simultaneous” and “sequential” lineups) will aid progress but it won’t finally answer the challenge.

It is axiomatic in high-reliability organizations that optimizing individual components is a poor route to overall system quality.²⁸ The double-blind sequential-lineup protocol is a more conservative screening test for guilt, but it isn’t a perfect one.²⁹ Individual cases with idiosyncratic histories will still have to be decided. Even after modernizing reforms, judges will still have to answer the question that medicine asks when offered a more conservative screening test for, say, prostate cancer or breast cancer: What does *the rest* of our system do with this new pattern of test results?

Could a “friendly and energetic alliance” of science and legal procedure give us new tools to “screen out” the higher number of cases that less conservative show-ups or traditional simultaneous lineups currently “screen in?” Or give us alternative ways to apprehend the perpetrators “missed” in the new, more conservative sequential lineups? To help judges gauge the impact of minor variations from accepted “best practice”? To develop a “forward-looking accountability” that helps us understand past mistakes to prevent future ones?

The answer to all of these questions will be “no” unless the judiciary plays an informed part. If eyewitness science does advance Wigmore’s “noble cause of justice,” it won’t happen in a single clap of thunder; it will happen as working judges apply the science with delicacy, to small details, in many decisions, and throughout the lives of many cases.

PRODUCERS AND INSPECTORS

Detectives speak of making cases; lawyers speak of trying them. The police operate a production stage in which they construct the case; the lawyers are elements of an inspection stage, during which the legal system evaluates the investigators’ product.

Judges can have an important impact on improving both the production stage and the inspection stage if they master the

basics of the eyewitness science. Something like that happened in medicine. When medical reformers accepted the “organizational accident” model of “iatrogenic” (caused by doctors or treatment) injuries to patients and understood that they were system errors, and not just the work of “bad apples,” they opened a window both on a more comprehensive understanding of past events and a more productive way to move forward as a profession to prevent future tragedies. Wrongful convictions are “iatrogenic” too, and judges can do something about them.

Direct judicial intervention in the business of producing evidence in eyewitness cases dates from at least the Warren Court’s exclusionary-rule cases in the 1970s. As several contributors to this issue point out, the scientific findings of recent decades have substantially undermined the Warren’s Court’s analysis of the problem. The sort of conscious police misconduct that can be deterred by exclusion is not the predominate issue, and the “reliability” test that the Warren Court instituted is largely obsolete. A modern approach to “best practices” in collecting eyewitness-memory evidence is plainly called for, and to their credit the law-enforcement authorities that must execute any best practices are moving to use science-based principles to renovate their procedures.³⁰

In this new context, exclusive reliance on the “nuclear option” of complete suppression of identification testimony every time some investigator varies marginally from the new “best practices” will be unworkable. This doesn’t mean that mistakes are inconsequential, but it does seem clear that judges will only rarely face one simple “in/out” decision about eyewitness testimony,³¹ while they will frequently (often many times within the same case) face smaller opportunities to exercise discretion about the admissibility of elements of testimony, the control of experts, the drafting of limiting instructions, and the provision of cautionary instructions,³² to deal with variances from the new accepted practices. Judges’ careful, graduated responses to the impacts of suboptimal practices will become crucial to their supervision of the production phase of cases.³³

The accuracy of these responses will depend on the individual judge’s knowledge of the basics of the science of memory, not on the judge’s mastery of broad lines of precedential appellate authority.³⁴ It is important, to take one example, that judges understand that the “strength” of a memory is a crucial factor in calculating the harm likely to have been caused by a suboptimal investigative practice. A “strong memory” formed in a lengthy encounter in bright light in calm conditions will be less affected by later procedural shortcomings than a

27. *Id.*

28. Donald Berwick, *Continuous Improvement as an Ideal in Healthcare*, 320 NEW ENGL. J. MED. 53, 53-55 (1989).

29. GARY L. WELLS, NANCY K. STEBLAY, & JENNIFER E. DYSART, A TEST OF THE SIMULTANEOUS VS. SEQUENTIAL LINEUP METHODS (2011).

30. See, e.g., Susan Gaertner & John Harrington, *Successful Eyewitness Identification Reform: Ramsey County’s Blind Sequential Lineup Protocol*, POLICE CHIEF, Apr. 2009, at 26; Amy Klobuchar et al., *Improving Eyewitness Identifications: Hennepin County’s Blind Sequential Lineup Pilot Project*, 4 CARDOZO PUB. L., POL’Y, & ETHICS J. 381 (2006).

31. This is qualified by the fact that a better understanding of the fundamental nature of eyewitness memory should make it clear to

judges that the idea of an “independent source” for courtroom testimony subsequent to a biased pretrial identification comes close to being a convenient fiction. See, e.g., Brandon Garrett, *Eyewitnesses and Exclusion*, 65 VANDERBILT L. REV. 201 (2012).

32. See Brian H. Bornstein & Joseph A. Hamm, *Jury Instructions on Witness Identification* 48 CT. REV. 22 (2012) (this issue).

33. See generally Richard A. Wise & Martin A. Safer, *A Method for Analyzing the Accuracy of Eyewitness Testimony in Criminal Cases*, 48 CT. REV. 22, 24-34 (2012) (this issue). For examples of courts wrestling with this recognition, see *Henderson*, 27 A.3d 872 (N.J. 2011); *Commonwealth v. Silva-Santiago*, 453 Mass. 782 (2009).

34. Smalarz and Wells, *supra* note 17.

“weak” memory formed in a fleeting, violent episode. But it is also crucial that the judges making assessments understand the sources of “strength of memory” and remember that “strength of memory” is not the same as “witness confidence.” Often, witnesses’ self-reports of “strength” indicate only memory contamination, not meaningful memory “strength.”³⁵

The final inspection stage of the criminal process—the jury trial—does address the diagnostic problem that Gary Wells emphasized in his path-breaking “system-variable” article: the riddle of how to *combine* the psychological factors present in an event and investigation that impact eyewitness reliability. The trial uses an ancient but flexible aggregating device: narrative. Jurors do not count and weigh piles of factors, or apply Bayesian formulae to arrive at probabilities; they generate and assess stories.³⁶ In the minds of the jurors, the psychological factors interact over time as a narrative unfolds. This feature of our inspection stage also has a fundamental political importance: the lay-citizen jury’s one-time concentration on a specific unique narrative provides a bracing challenge to the official practitioners’ endemic tendency to believe that since we know the odds in our fields we can simply play those odds. The professionals tend to believe that if we know what happens 90% (or 80%, or even 51%) of the time, then we know what to do 100% of the time. If things go right under the story model, every accused gets an individualized jury judgment, not a roll of the probabilistic dice.

An important part of the trial judge’s role is to manage the “story-model” core of the jurors’ work, and the science of identification indicates that eyewitness cases present particularly difficult problems in this regard. This task doesn’t require a Ph.D. in psychology, but it does require more than reading appellate-suppression and expert-testimony precedents.

Many jurors, if left to their own devices, will default to a “videotape” story—the witness recorded the event like a camera, stored it on a permanent tape, and is now replaying it—that is contradicted by the scientific truth that memory evidence is malleable “trace evidence.” It is also pretty clear that traditional tools such as cross-examination will be insufficient to convey much of the new science of memory because the jurors’ vulnerability is not on the level of specific missing pieces of data (e.g., “the witness was/was not confident”) but on the level of the general background interpretive principles that no cross-examiner can reach (e.g., “confidence means accuracy”) no matter how clever his or her questions.³⁷

The “estimator variable” story of the crime event must be complemented by the “system variable” story of the *investigation* before the story-model inspection can be effective. The eyewitness research indicates that in administering the story model, judges will have to attend to not only general juror “common-sense” principles that may be mistaken, but also specific pieces of data that scientists have learned are necessary to the story-testing process but that upstream operators have

not preserved or disclosed. These data will not be available unless science-informed judges act to make them available.

This means that judges must incorporate into their daily practice the recognition that the production and the inspection stages of an eyewitness-based prosecution are reciprocally related. Inevitably, while the judges “downstream” are trying to adjust for the exigencies of upstream investigative operations, the “upstream” law-enforcement operators are trying to adjust their conduct in anticipation of the inspection that awaits their cases downstream.

It is axiomatic in medicine and other industries that end-of-process inspection schemes, although necessary components of their systems, are poor routes to overall system quality.³⁸ Practitioners who are subject to inspection are resourceful in both avoiding the inspection altogether or in gaming the inspection when they cannot avoid it. Those being inspected usually end up owning the process, and their primary goal is usually their own safety. Criminal-justice-system operators are not immune to these tendencies. The fact that only a tiny portion of criminal cases receives jury scrutiny certainly has something to do with the costs of jury trials in terms of time and money, but it also reflects professional practitioners’ disinclination to submit to inspection by unpredictable lay jurors, especially when that inspection takes place in an exposed zero-sum courtroom contest where one side wins (and one side loses) everything.

Here’s an example. There is a segment of the eyewitness-exoneration list that catalogs trial prosecutors’ failures to turn over exculpatory material. It does not show that those prosecutors lusted to frame known innocents, but rather it illuminates an impulse to shape the adversary trial inspection stage so that it comes out (from the prosecutors’ perspective) the “right” way. Sometimes, prosecutors don’t disclose eyewitness exculpatory material because they simply don’t understand what factors are influential in eyewitness performance. Sometimes, prosecutors withhold information to convict the men the prosecutors believe are guilty without interference from “red herrings” that defense lawyers might manufacture out of dissonant facts.

The trial prosecutors in the wrongful-conviction *Brady* cases, like workers in most production processes, evidently adopted a “covert work system.”³⁹ They decided to evade formal disclosure requirements and buried alternative narratives because they believed sharing the exculpatory facts would interfere with achieving what they saw as the “real” goal tac-

“[J]udges must incorporate into their daily practice... the production and the inspection stages of an eyewitness-based prosecution....”

35. *Id.* at 18.

36. See Nancy Pennington & Reid Hastie, *Explaining the Evidence: Tests of the Story Model for Juror Decision Making*, 62 J. PERSONALITY & SOC. PSYCHOL. 189 (1992).

37. DOYLE, *supra* note 2, at 35-49. Jules Epstein, *The Great Engine That Couldn't: Science, Mistaken Identifications, and the Limits of Cross-*

Examination, 36 STETSON L. REV. 727 (2007).

38. ATUL GAWANDE, *THE CHECKLIST MANIFESTO: HOW TO GET THINGS RIGHT*, 185 (2009).

39. See, e.g., *Connick v. Thompson*, 131 S. Ct. 1350 (2011); David D. Woods, *Conflicts Between Learning and Accountability in Patient Safety*, 54 DEPAUL L. REV. 485 (2005).

“Judges cannot dictate all the choices made by the system’s other actors, but they can influence them.”

itly assigned to them by officials (or the public) to whom they were accountable. Turning a blind eye to these practices encourages upstream tunnel vision by rewarding practitioners’ surrender to tunnel vision with a “cleaner” trial inspection for the hypothesis that they prematurely decided is accurate.

Tunnel vision is a “cause” of wrongful convictions, but tunnel vision is also an *effect* of the sharp-end operators’ discomfort with the demands of the end-stage inspection machinery. A resulting wrongful conviction is an “organizational accident”: the police make the wrong choice; the prosecutors buy it too quickly; and the defense and the jury are crippled in their inspectors’ roles.

One of the lessons of the eyewitness-exoneration cases is that judges must develop (and incorporate in their inspection-stage calculations) an awareness of the gravitational pull away from comprehensive and transparent investigation that is always acting on production-stage practitioners. Science-conscious judges can put a brake on this rush down the “organizational-accident” tunnel by making it clear that *they* know what matters in eyewitness-evidence collection and that they will insist on detailed documentation and disclosure. The story model of aggregating eyewitness factors cannot work if details (e.g., confidence-boosting comments, exposure to co-witnesses, neglected alternative suspects) are not available to be considered as part of the story. Diagnosing eyewitness errors requires weighing not just catastrophic contradictions (e.g., the defendant is tall, the crime-night police report described a midget) but also small narrative details (e.g., brief exposures to co-witness accounts, or mug-book pictures of the defendant) that accumulate and ultimately constitute the story of inadvertently corrupted eyewitness memory traces.

The categorical exclusion of identification evidence because of misconduct may become less frequent as law enforcement gradually absorbs and adapts the modern “system variable” science. But pretrial hearings that will allow the trial judge to assess (on some basis other than laconic police reports) the source and quality of the eyewitness evidence that is *not* excluded and to decide which judicial tools—for example, *in limine* edits of evidence, cautionary instructions—will assist the jurors’ story-model inspection and will become more important.⁴⁰ Unless alert and informed judges play an active role in protecting these aspects of story-model testing, sharp-end practitioners worried about inspections will simply shift from “don’t turn it over” to “don’t write it down,” a practice that will end up hampering not only inspectors, but their fellow investigator-producers, who could be exploring alternative theories and correcting their tunnel vision.

JUDICIAL-SYSTEM LEADERS: BEYOND INSPECTION

There is no arrangement of gears and switches in criminal justice, no system in that sense that we can reach for and fix with a wrench or a hammer. But, like it or not, the world of criminal justice is a complex functioning ecosystem like a pond or a swamp where well-meaning actions on this coast can have disastrous, unanticipated impacts on the far shore. Ignoring this fact will fulfill the axiom that the cause of problems is solutions. Judges cannot dictate all the choices made by the system’s other actors, but they can influence them. In fact, the nature of the system guarantees that judges cannot *avoid* influencing those choices. Even judicial silence and inaction will always have an impact.

There is opportunity as well as danger in this interdependency of criminal justice’s operators. A recent episode in the history of the “friendly and energetic alliance” provides an example. Law-enforcement practitioners were intrigued in the aftermath of the DNA exonerations by the potential of the “double-blind sequential” system-variable approach, but they were uncomfortable that it had not been tested in the field. A well-meaning, go-it-alone attempt by the general counsel of the Chicago Police Department to conduct a field study to fill the gap resulted in a kind of scientific travesty.⁴¹ But when an actual alliance of science and law enforcement was formed by a team composed of researchers, the Police Foundation, the Center for Problem-Solving Policing, and the American Judicature Society to design and execute a scientifically rigorous field examination of the issue, it largely vindicated the hopes of the advocates of that reform.

In the process of organizing the study the researchers developed—and the frontline practitioners tested the practicality of—a laptop-housed program that allows for both the effective administration and the meticulous documentation of double-blind sequential eyewitness-identification procedures. Seen from the system level, this is an example of errors spurring us to learn how the conditions facing the sharp-end investigators *and* the inspecting trial courts could both be substantially improved by an investment made by officials distant from the scene in cooperatively identifying and disseminating a relatively simple (and relatively inexpensive) technological improvement. As we enter an era in which every patrol car will have a laptop and every court will face subtle eyewitness evidentiary issues, this is a development that all of the operators jointly responsible for eyewitness “organizational accidents” can work together to accelerate. Recognizing that the judiciary doesn’t draft law-enforcement budgets or vote on law-enforcement appropriations isn’t quite the same thing as saying that the judiciary can’t find ways to signal its support for such an effort.

But it is also worth focusing for a moment on the *practice* of nonblaming learning from error, apart from that practice’s immediate products.

Working steadily on “organizational-accident” error analysis can create an increased system consciousness among the practitioners who staff the components of the criminal process.

40. See, e.g., *State v. Henderson*, 27 A.3d 872 (N.J. 2011).

41. Daniel Schacter et al., *Policy Forum: Studying Eyewitness Identifications in the Field*, 32 L. & HUM. BEHAV. 3, 4 (2007).

A disciplined commitment to non-blaming, team analysis of error can lay the foundation for mobilizing the new ideal of continuous quality improvement that is transforming the culture of contemporary medicine in criminal justice.

Inspection of the prosecution's case during an adversary trial before a lay jury is a permanent feature of our system. It expresses fundamental American convictions about the relationship between the accused individual and the state. But the goal of the trial process is to protect *this* innocent citizen from the state. The DNA exonerations have raised concerns about the adversary trial's weaknesses even in that specific role,⁴² but no one ever claimed that the trial's role is to analyze the investigative and charging processes and make them more reliable in future cases. A jury that believes that it has caught a faulty investigation says "not guilty" and nothing more. Appellate courts review the legal procedures; they do not reconsider the facts, and their review is entirely backward looking. Both are necessarily uninformative.

The criminal justice system currently lacks the capacity for "forward-looking accountability"⁴³ that not only catches past mistakes, but also anticipates and precludes future ones.

The challenge for the judiciary presented by a new "organizational-accident" understanding of how eyewitness errors happen is not protecting a presumptively safe system from the misconduct of sloppy (or even evil) human components—the approach taken by the Warren Court in its misconduct-based suppression cases. The challenge judges will confront is how to invigorate and support a culture of constant, routine attention to safety and reliability in the criminal process.

The missing weapon in our approach to error is not the once-in-a-decade, blue-ribbon panel of dignitaries at the chief justice and superintendent level, convened to redesign the architecture of the criminal justice system. We have examples of that vehicle now, and the judiciary has played a leading role in several of them.⁴⁴ When the goal is changing structural elements of the system by legislation or rulemaking, the political heft of those high-ranking players can be useful, even essential.⁴⁵

What we are missing is a consistent commitment to regular, routine review of known errors and "near misses," conducted by experienced practitioners and stakeholders (for example, victims' rights professionals) supplemented where appropriate by subject-matter experts and (at least in the beginning) by specialists in analyzing the sources of system error and in the error-review process itself. As Lucien Leape argued in his seminal essay *Error in Medicine*:

The emphasis is on *routine*. Only when error is accepted as an inevitable, although manageable, part of everyday practice will it be possible to shift from a punitive to a creative frame of mind that seeks out and identifies the underlying system failures.⁴⁶

For many reasons the best hope for breathing life into the "friendly and energetic alliance in the noble cause of justice" may lie in the judiciary: in judges who exercise their power to convene criminal justice stakeholders outside their familiar adversary bunkers. The alliance can serve the noble cause not only by asking the system's actors to do a better job playing "Whac-A-Mole"

and catching past errors one at a time, but also by asking them to uncover and address the abiding latent weaknesses of the system that will survive to cause future errors.

What if, when the next wrongful eyewitness conviction is revealed, the local judiciary amazes the world by calling for a dispassionate, all-stakeholders examination of the error? Or what if, when DNA results come back from the lab six months after an arrest and show that law enforcement arrested the wrong guy on the night of the crime, the judges suggest that a team examination of this "near miss" might pay dividends, both in terms of what worked and what nearly didn't?

Just as all aviation-industry participants and the public expect the National Transportation Safety Board to convene a mixed team of specialists to give an account of what happened when a plane goes down, criminal practitioners and the public could learn to expect that we will marshal a team including an investigator or patrol supervisor, a prosecutor, a forensic scientist, a defender, a judge, a victims' representative, and the jurisdiction's risk management officers, joined by additional specialists as needed, in a nonblaming process of dissecting the record of what happened and sharing the account they have developed. The goal would be to understand the gritty facts, to do the sort of clinical fact-finding that inevitably suffers when everyone in a turf-conscious, blue-ribbon group is anxiously looking over his or her shoulder at potentially sweeping and unwelcome law reforms.

Continually working on improving system reliability means changing the system's culture, not just its architecture. Overhauling institutional arrangements, identifying best lineup practices, and devising checklists, as difficult as these tasks might be, are the easy parts. Working on changing the culture means concentrating on giving a primary place to workmanship and professionalism instead of blame and discipline. It means learning—as medicine learned—to treat errors as "sentinel events" to be studied, not as embarrassments to be buried.

The history of the eyewitness cases illuminates the potential in a coherent program of nonblaming learning from error that

"Working steadily on 'organizational-accident' error analysis can create an increased system consciousness...."

42. Dan Simon, *The Limited Diagnosticity of Criminal Trials*, 64 VANDERBILT L. REV. 143 (2011).

43. Virginia A. Sharpe, *Promoting Patient Safety: An Ethical Basis for Policy Deliberations*, HASTINGS CENTER REPORT, SPECIAL SUPPLEMENT NO. 33 (2003).

44. See Christine C. Mumma, *The North Carolina Actual Innocence*

Commission: Uncommon Perspectives Joined by a Common Cause, 52 DRAKE L. REV. 647 (2004).

45. Stephen Saloom, *Adversaries as Allies: Joining Together to Prevent Criminal Injustice*, 73 ALBANY L. REV. 1235 (2010).

46. Lucian L. Leape, *Error in Medicine*, 272 JAMA 1851, 1854 (1994).

includes the evaluation of “near misses,” and offers rewards both within local systems and across scattered systems. A common national template for error review, enacted locally and informed and challenged by diverse local experiences, could substantially mitigate the fragmentation of American criminal justice.

These advantages can be multiplied if a simple mechanism—a clearinghouse, or a wiki-style community of practitioners, researchers, and policymakers—could be developed for distributing and commenting on the reports of errors.⁴⁷ Reading of a distant system’s experience of completed accidents can alert currently isolated practitioners to the operation of dangerous latent features in their own local systems. Reading studies of remote “near misses” can reveal both those dangerous latent features and potential fail-safe devices or procedures that are *not* present locally, but which provided resilience and kept the near miss in another jurisdiction from becoming a tragic “hit.” It can counteract the tendency of today’s best practice to calcify into a ceiling that blocks future improvements.

After an exoneration it is often very easy to see in hindsight where a wrong decision was made. But congratulating ourselves on recognizing past bad choices won’t get us very far. We have to learn why the last bad decision looked like a *good* decision from the perspective of the mistaken detective or prosecutor or defender or judge at the time it was made. If we don’t, the root causes of the last tragedy will continue to lie in wait for the next decision maker who comes along. Accounts of eyewitness wrongful-conviction cases give striking evidence of how much we could learn about latent system defects from a close, all-stakeholders analysis that incorporates the scientific contributions⁴⁸ that follow in this issue and the operations-oriented insights of the sharp-end participants who do the work on the streets and in the courts.

The judiciary is uniquely well placed to stake out the common ground on which criminal-justice-system actors could meet, to invite the participants onto that ground, and to help them to defend that ground against the short-term pressures for public pillories filled with scapegoats.

The DNA exonerations have killed the illusion of an infallible justice system forever. From now on, the legitimacy of the criminal justice system in the public’s eye will depend significantly on that system’s willingness to confront its own failures.⁴⁹ We will never have an exact count of those failures, but when the most careful analyses we can muster suggest that the wrongful-conviction rate may be as high as 6-15% in sexual-assault cases,⁵⁰ the exact count becomes almost irrelevant. All of us in criminal justice have some explaining to do, and we could start by explaining our practices to each other, without trying to point fingers and assign blame.

We have some prevention to do as well. When medicine

adopted its new approach to iatrogenic “sentinel events” and moved toward self-consciously creating a culture of safety, it quickly saved 120,000 patients’ lives in eighteen months.⁵¹ The eyewitness cases, with their wrongfully convicted defendants and their wrongfully free perpetrators (and the later victims those perpetrators find) make a strong argument that the criminal justice system’s natural leaders—the judges—armed with an important body of scientific knowledge available for application, could do some leading in that direction.



James M. Doyle, a Visiting Fellow at the National Institute of Justice (NIJ) during 2012, is a veteran litigator and writer. He is the author of True Witness, the history of the collision between the science of memory and the legal system and the co-author (with Elizabeth Loftus) of Eyewitness Testimony: Civil and Criminal, a treatise for lawyers in eyewitness cases. He has published numerous articles on evidence, race in criminal justice, and capital punishment. He served as the founding director of The Center for Modern Forensic Practice at the John Jay College of Criminal Justice. The former head of the statewide Public Defender Division of the Committee for Public Counsel Services in Massachusetts, he practices law in Boston as counsel to the firm of Carney & Bassil. He received his B.A. from Trinity College, in 1972, J.D. from Northwestern University School of Law in 1975, and LL.M from Georgetown University Law Center, where he was an E. Barrett Prettyman/LEAA Fellow, in 1979. He expresses his thanks to NIJ and its director, Dr. John Laub, for supporting his current NIJ Fellowship and notes the opinions in this article are not necessarily the views of NIJ. Email: lamesdoyle@gmail.com

47. See Doyle, *supra* note 24.

48. An excellent comprehensive analysis of the scientific literature bearing on the criminal process as a whole is found at DAN SIMON, IN DOUBT (2012).

49. See generally TOM R. TYLER, WHY PEOPLE OBEY THE LAW: PROCEDURAL JUSTICE, LEGITIMACY AND COMPLIANCE (2006).

50. JOHN ROMAN, KELLY WALSH, PAMELA LACHMAN & JENNIFER YAHNER, POST-CONVICTION DNA TESTING AND WRONGFUL CONVICTION 57 (2012).

51. CHARLES KENNEY, THE BEST PRACTICE: HOW THE NEW QUALITY MOVEMENT IS TRANSFORMING MEDICINE 270-272 (2008).