Dædalus

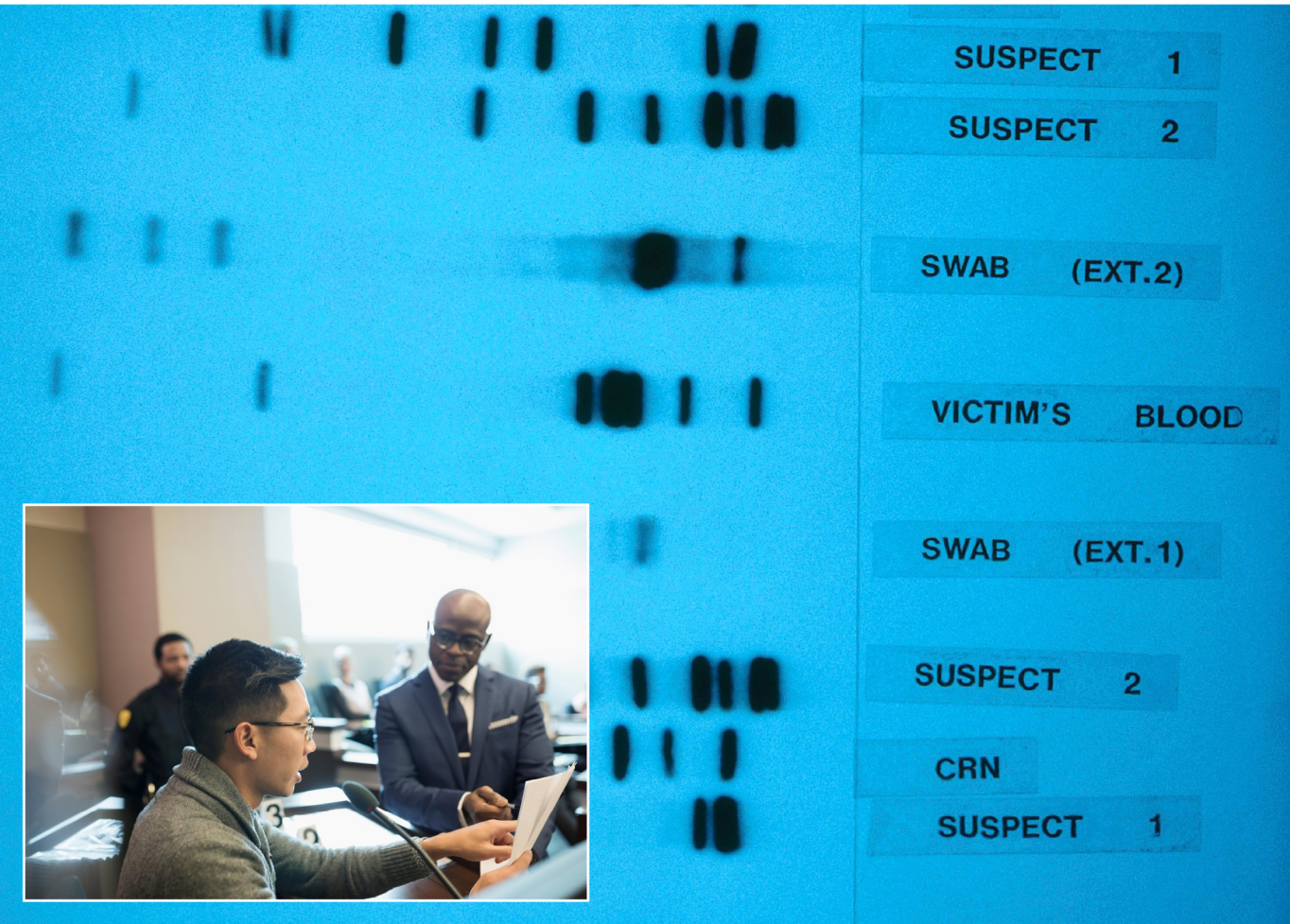
Journal of the American Academy of Arts & Sciences

Fall 2018

Science & the Legal System

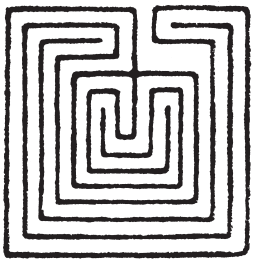
Shari Seidman Diamond & Richard O. Lempert,
guest editors

with Sheila Jasanoff · Linda Greenhouse
Jules Lobel · Huda Akil · Rebecca S. Eisenberg
Robert Cook-Deegan · Jed S. Rakoff
Elizabeth F. Loftus · Jennifer L. Mnookin
Joseph B. Kadane · Jonathan J. Koehler
Nancy Gertner · Joseph Sanders · Daniel L. Rubinfeld
Joe S. Cecil · Valerie P. Hans · Michael J. Saks
David Baltimore · David S. Tatel · Anne-Marie Mazza


SUSPECT 1
SUSPECT 2
SWAB (EXT.2)
VICTIM'S BLOOD
SWAB (EXT.1)
SUSPECT 2
CRN
SUSPECT 1

Dædalus

Journal of the American Academy of Arts & Sciences



"Science & the Legal System"

Volume 147, Number 4; Fall 2018

Shari Seidman Diamond & Richard O. Lempert, Guest Editors

Phyllis S. Bendell, Managing Editor and Director of Publications

Peter Walton, Associate Editor

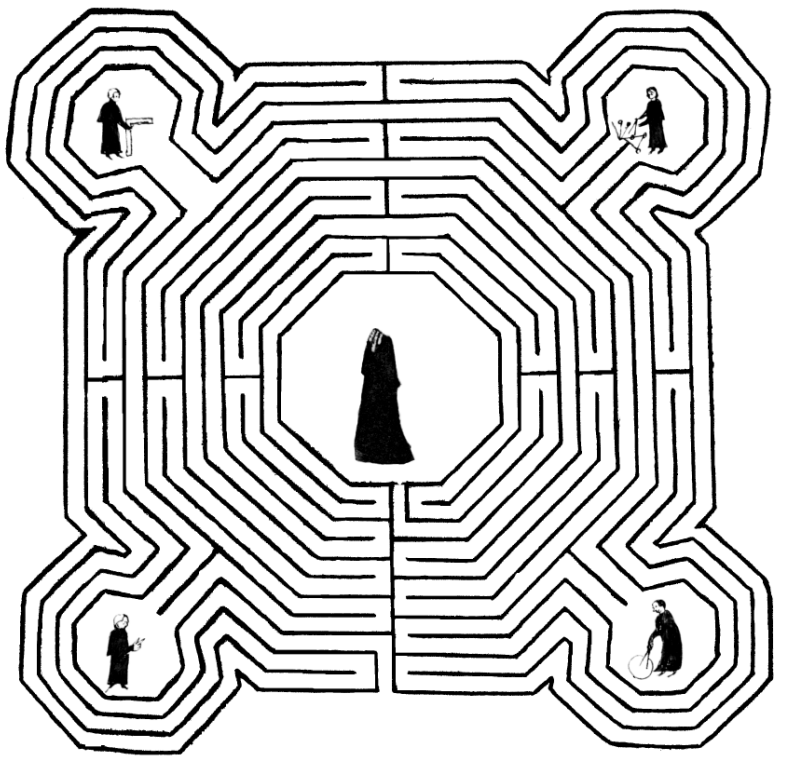
Heather M. Struntz, Senior Editorial Associate

Inside front cover: When scientists appear in court today, they are not being prosecuted for their scientific views, but what science tells them may be ignored or, worse, faulty analyses may prove more convincing than sound science. The challenge of getting courts to recognize and accept sound science is still with us. (Top) Joseph-Nicolas Robert-Fleury, *Galileo before the Holy Office*, 1847, Paris, Musée du Luxembourg. (Bottom) DNA sequencing gel © 2018 by Larry Mulvehill; witness testimony © 2018 by Hero Images Inc.

Contents

Dædalus

Journal of the American Academy of Arts & Sciences



The pavement labyrinth once in the nave of Reims Cathedral (1240), in a drawing, with figures of the architects, by Jacques Cellier (c. 1550–1620)

Dædalus was founded in 1955 and established as a quarterly in 1958. The journal's namesake was renowned in ancient Greece as an inventor, scientist, and unriddler of riddles. Its emblem, a maze seen from above, symbolizes the aspiration of its founders to "lift each of us above his cell in the labyrinth of learning in order that he may see the entire structure as if from above, where each separate part loses its comfortable separateness."

The American Academy of Arts & Sciences, like its journal, brings together distinguished individuals from every field of human endeavor. It was chartered in 1780 as a forum "to cultivate every art and science which may tend to advance the interest, honour, dignity, and happiness of a free, independent, and virtuous people." Now in its third century, the Academy, with its more than five thousand members, continues to provide intellectual leadership to meet the critical challenges facing our world.

Dædalus Fall 2018
Issued as Volume 147, Number 4

Editorial offices: *Dædalus*, American Academy of Arts & Sciences, 136 Irving Street, Cambridge MA 02138. Phone: 617 576 5085. Fax: 617 576 5088. Email: daedalus@amacad.org.

Library of Congress Catalog No. 12-30299.

Dædalus (ISSN 0011-5266; E-ISSN 1548-6192) is published quarterly (winter, spring, summer, fall) by The MIT Press, One Rogers Street, Cambridge MA 02142-1209, for the American Academy of Arts & Sciences. An electronic full-text version of *Dædalus* is available from The MIT Press. Subscription and address changes should be addressed to MIT Press Journals Customer Service, One Rogers Street, Cambridge MA 02142-1209. Phone: 617 253 2889; U.S./Canada 800 207 8354. Fax: 617 577 1545. Email: journals-cs@mit.edu.

The typeface is Cycles, designed by Sumner Stone at the Stone Type Foundry of Guinda CA. Each size of Cycles has been separately designed in the tradition of metal types.

Subscription rates: Electronic only for nonmember individuals – $51; institutions – $168. Canadians add 5% GST. Print and electronic for nonmember individuals – $57; institutions – $210. Canadians add 5% GST. Outside the United States and Canada add $24 for postage and handling. Prices subject to change without notice. Institutional subscriptions are on a volume-year basis. All other subscriptions begin with the next available issue.

Single issues: $15 for individuals; $38 for institutions. Outside the United States and Canada, add $6 per issue for postage and handling. Prices subject to change without notice.

Claims for missing issues will be honored free of charge if made within three months of the publication date of the issue. Claims may be submitted to journals-cs@mit.edu. Members of the American Academy please direct all questions and claims to daedalus@amacad.org.

Advertising and mailing-list inquiries may be addressed to Marketing Department, MIT Press Journals, One Rogers Street, Cambridge MA 02142-1209. Phone: 617 253 2866. Fax: 617 253 1709. Email: journals-info@mit.edu.

Printed in the United States by The Sheridan Press, 450 Fame Avenue, Hanover PA 17331.

Newsstand distribution by Ingram Periodicals Inc., 18 Ingram Blvd., La Vergne TN 37086.

Postmaster: Send address changes to *Dædalus*, One Rogers Street, Cambridge MA 02142-1209. Periodicals postage paid at Boston MA and at additional mailing offices.

Introduction

Shari Seidman Diamond & Richard O. Lempert

Experts bedeviled the legal system long before seventeenth-century Salem, when the town's good citizens relied on youthful accusers and witchcraft experts to identify the devil's servants in their midst. As in Salem, claims of expertise have often been questioned and objections raised about the bases of expert knowledge. Expertise, then and now, did not have to be based on science; but the importance of science and the testimony of scientific experts has since medieval times been woven into the fabric of the English jurisprudence that Americans inherited. In cases as long ago as 1299 we find examples of courts seeking help from "scientists." In that year, physicians and surgeons in London were called on to advise the court on the medical value of the flesh of wolves.[1] In 1619, two physicians offered the opinion that a wife could bear a legitimate child "forty weeks and nine days" after the death of her husband.[2] Throughout this period, medical authority was called on by the coroners' courts to determine whether a death was due to suicide or to other causes, a crucial determination because suicide was a felony that entitled the Crown to take possession of a deceased's estate.[3] Medical testimony is still the most common form of scientific expertise presented in court, but expert advice on legal matters has expanded exponentially, reflecting the enormous range of scientific knowledge that modern scholarship has produced.

Although recognizing the need for scientific assistance, judges soon learned that sources claiming scientific expertise did not always agree. For example,

SHARI SEIDMAN DIAMOND, a Fellow of the American Academy since 2012, is the Howard J. Trienens Professor of Law and Professor of Psychology at the Northwestern University Pritzker School of Law, and Research Professor at the American Bar Foundation.

RICHARD O. LEMPERT, a Fellow of the American Academy since 1993, is the Eric Stein Distinguished University Professor of Law and Sociology, emeritus, at the University of Michigan.

(Complete author biographies appear at the end of the essay.)

doi:10.1162/DAED_x_00516

in the 1781 trial of *Folkes v. Chadd*, the issue was whether the construction of an embankment, as opposed to natural forces, had caused the deterioration of Wells Harbor. The first trial introduced engineering testimony from a well-credentialed Fellow of the Royal Society. By the third trial in 1783, prestigious engineering experts testified on both sides and were subjected to vigorous cross-examination. The disagreement, in retrospect, was understandable: more than two hundred years later, science still cannot provide a definitive answer to the question posed in that litigation.[4] Yet the legal system then as now needed to resolve the dispute between the parties, and the scientific evidence offered was the best they had to work with. As the trial system and the law of evidence developed, courts and juries have continued to struggle to make use of the conflicting expert advice they receive. Judges and juries, lacking the scientific knowledge of experts, both face difficult challenges in understanding and applying expert scientific testimony. Not surprisingly, they occasionally get the science they are supposed to evaluate wrong, and what the legal system has accepted as sound science has not always withstood the test of time.

How well factfinders do in understanding and applying science is a matter of some controversy, but it is not the only issue that arises at the interface of law and science. The two fields are in many ways culturally distinct. Good science often involves the withholding of judgment until more evidence has accumulated. The law requires that decisions be reached upon the conclusion of trials regardless of gaps in the available evidence. Science seeks empirical truths regardless of their implications, and scientists ideally share in a common truth-seeking mission. Litigants aim at persuading a judge or jury to favor their side regardless of where the truth lies; harsh questioning and emotional appeals are not out of bounds if they serve that end, even when it is scientists being questioned. Often in modern litigation, the law must be informed by scientific evidence as communicated by the views of the scientists who present it. These are typically experts chosen and paid by parties because, regardless of the law's needs, scientists, with rare exceptions, cannot be forced to contribute what they know. Science is in principle always open to revision as additional evidence accumulates. The law can be slow to change and its treatment of science may be determined by precedent, even when a scientific consensus recognizes that the science that supported the precedent is no longer regarded as sound.

The essays in this volume deal with tensions and areas of overlapping interest at the interface of science and the legal system. Many of the essays are written by scientist-lawyer teams. This is no accident; in selecting authors we tried wherever possible to match across disciplines to highlight and bridge potential gaps in perspectives. In some cases, we selected single authors who themselves are both scientists and legal scholars. Our goal was to avoid the silo mentality that too often creates obstacles to useful discourse between science and law.

The essays in this issue are divided into three sections. The essays in the first section examine the science-law interface by focusing attention on two sets of key players: the judges who determine what scientific evidence will be considered by the legal system, and the scientists and engineers with the expertise to provide that assistance. The authors of the first two essays have closely studied the history, discourse, and decision-making of U.S. courts when they are called on to deal with scientific evidence as gatekeepers and decision makers. The third essay provides a perspective from the other side of the law-science divide. It presents the first published survey

results from a sample of distinguished scientific and engineering experts who were asked about their views of the legal system and about their participation in it (or not).

The five essays in the second section provide insights into the interactions between scientific expertise and the legal system by focusing on specific fields: neuroscience, patents, eyewitness identification, forensic evidence as a whole, and fingerprint evidence in particular. Each of these contributions highlights what science can offer, but also analyzes the obstacles that arise in obtaining and evaluating scientific advice in a legal context.

The authors in the third section tackle the difficult procedural challenges posed by the interaction between scientific experts and legal factfinders. These three essays consider modest and not-so-modest changes to the traditional conduct of American legal proceedings that might improve both the presentation and evaluation of scientific evidence.

The issue closes with a look at the continuing dialogue between members of the scientific and legal communities.

Now for a closer look.

In the volume's opening essay, Sheila Jasanoff addresses an issue fundamental to any discussion of science and the law: what determines the reception given ostensibly scientific claims when they enter the legal system and are reinterpreted in a legal context? Jasanoff argues that judicial common sense, rooted in judges' cultural understandings, forms the lens through which scientific claims are assessed by courts. She makes a powerful case for her view of how judicial authority and judges' commonsense understandings of the import and validity of scientific claims provide the standards that effectively determine how scientific evidence is perceived and used by courts. Her perspective cautions against analyses that too frequently begin and end with *Daubert v. Merrell Dow Pharmaceuticals*, the Supreme Court case that firmly established the judge's role as gatekeeper when courts are offered scientific evidence. She uses an extensive analysis of *Kumho Tire Co. v. Carmichael*, a case that made it clear that *Daubert* extended to engineering and technical experts to show how the standards for admitting scientific evidence, which the *Daubert* court tried to draw from their understanding of how scientific truths are established, are easily submerged by judges' commonsense perspectives on what methods and theories make for sound scientific or technical conclusions. Her analyses of later cases highlight limits on the guidance that *Daubert* can give, for science may background some legal questions but be unable to answer them.

In closing her essay, Jasanoff argues that one cannot expect judges to think like scientists when evaluating scientific evidence, but she contends that we can demand of judges who confront scientific issues more than unreflective common sense. The challenge is not to make scientists of judges but rather to reflect on how judges should go about thinking about science and to find ways of encouraging judges to appreciate what science can tell them and see beyond their own common sense. Although Jasanoff does not say it, the task becomes more difficult as ideology affects judgments.

Linda Greenhouse, closely scrutinizing how members of the U.S. Supreme Court have responded to scientific evidence, provides a detailed study of the ways that law and medical science have intertwined in the jurisprudence surrounding abortion, beginning with *Roe v. Wade*. Greenhouse tells us that the case law began with a focus more on protecting medical doctors in their exercise of professional judgment from the threat of prosecution than on the interests that pregnant women had in choosing to terminate a pregnancy. As Greenhouse de-

scribes the case law, an elaborate dance has been occurring between science and the law, with each in turn taking steps forward and back. Which partner is moving forward depends on legal understandings of the deference courts owe legislative fact-finding and limits on this deference when the facts do not even arguably stand up to scientific scrutiny.

In *Whole Woman's Health v. Hellerstedt*, for example, the question was whether courts should defer to the Texas legislature's assertion that protecting the safety of women getting an abortion requires that doctors who perform abortions must have hospital admitting privileges (a requirement that would, in effect, close most abortion clinics). The U.S. Supreme Court rejected the legislation, which ignored the compelling medical evidence that requiring hospital privileges does nothing to protect women needing more medical attention than a clinic can provide. But the path to the Supreme Court's decision was rocky. The decision of the District Court that initially heard the case, finding that the facts were inconsistent with the legislative claim, was reversed by the Circuit Court on appeal on the respectable-in-theory but unjustified-in-context claim that federal courts should defer to legislative fact-finding on the need for health-related regulation. The Court of Appeals also refused to stay its decision pending appeal to the Supreme Court. By the time the Supreme Court eventually upheld the District Court's decision enjoining enforcement of the statute, in 2016, about half of Texas's abortion providers had permanently closed their doors. Although science-based evidence eventually prevailed in this case, an important lesson from this dance between law and science is that judges vary in their openness to what science and technology can offer, with ideology sometimes motivating a failure to accept even strong scientific evidence.

We, Shari Diamond and Richard Lempert, coeditors of this volume, describe the results of a survey that many Academy members participated in – our thanks! Conducted with the cooperation of the American Academy of Arts and Sciences, the survey examines the views of the legal system held by some of the nation's most distinguished scientists and engineers, including what motivates them to participate or to refuse to participate in lawsuits when asked. We began the project with some doubt that the legal system was soliciting assistance from the kinds of scientific and engineering experts whose accomplishments have led to Academy membership – or that, perhaps, such experts were being asked but were unwilling to participate. The results showed that these concerns were unwarranted. A majority (54 percent) of respondents reported having been asked for advice, and most of those asked had agreed to participate at least once.

Nonetheless, we found that the experts reported that lack of time frequently limited their participation, and that they sometimes turned down requests due to a discrepancy between their area of expertise and the scientific issues they were asked about, suggesting that greater participation might be promoted through a more effective matching system. In addition, respondents endorsed several potential changes in procedures used by the legal system that might increase their willingness to participate. Some of these potential changes are discussed in greater depth in the third section of this volume. Finally, we found an intriguing relation between participation and belief in the ability of the legal system to deal well with scientific matters, including some evidence that participation fuels higher opinions. This is a relationship that deserves further investigation.

More than any other contribution to this volume, Jules Lobel and Huda Akil's essay

on law and neuroscience is positioned on an active and changing border between law and science. Courts are increasingly being asked to consider neuroscience evidence. To date, neuroscience has had the greatest impact on legal processes on the criminal side, where neuroscience evidence can reveal deficiencies in an accused's brain that suggest the intent behind a criminal action was in part the result of physiological abnormalities. The evidence can even have constitutional significance, as in *Roper v. Simmons*, the case that barred executing juveniles, influenced in part by evidence regarding the neurological development of youthful brains. Civil litigation too may be transformed by neuroscience. The civil justice system has long resisted awarding damages or other relief based on emotional pain unaccompanied by noticeable physical harm. Such suits were regarded with suspicion because of the subjective nature of claims of emotional harm and the difficulties of finding objective proof. But to the extent that neuroscience can provide imaging evidence that a claimant's brain deviates from normal human physiology, the claim of emotional harm is objectively supported and physical harm is shown to be present.

Much of the Lobel-Akil essay is devoted to a close look at cases arguing that long-term solitary confinement is unconstitutionally cruel and unusual. Although lawyers opposing extended solitary confinement have few if any scientifically rigorous studies of people to draw on, considerable animal research and a body of neuroscience theory supports the claim that people's brains undergo seriously harmful and likely permanent changes when they are denied social contact and environmental stimulation over long periods of time. To the extent this new research moves the dial on the practice and legality of long-term solitary confinement, it will also tell us something about the law. Most people, judges included, do not need neuroscience to convince them of the horror of isolating people in small confined spaces with almost no social contact for years on end. Yet the law may need scientific evidence in support of what almost everyone knows before it will discard the fiction that solitary confinement differs simply in degree, rather than in kind, from the normal deprivations that anyone imprisoned suffers. This may be one area in which scientific evidence can resolve differences between conflicting commonsense beliefs.

Rebecca Eisenberg and Robert Cook-Deegan write about an area in which science and the law are intertwined to the point where they cannot be untangled: the U.S. patent system. The authors focus their attention on the Bayh-Dole Act, which changed prior law by not only allowing but also encouraging organizations that develop patentable inventions through research funded by federal agencies to acquire proprietary rights to these inventions. The goal was to promote the commercialization of the fruits of federally funded science. Universities were the most visible intended beneficiaries, and the image of universities as entities working for the common good by advancing and sharing knowledge created halo effects without which Bayh-Dole might never have become law. The benefits of Bayh-Dole were, however, later extended from nonprofits and small businesses to large corporations by a low visibility amendment.

Eisenberg and Cook-Deegan document the effects of Bayh-Dole by focusing on how universities responded to their new rights in light of the income streams these rights enabled. In many cases, it appears, monetary concerns dwarfed whatever perceived commitment to the common good universities benefited from when the case was made for Bayh-Dole and in their later patent-related legislative lobbying. In a number of instances, universities claimed

patent rights to stifle or extract profits from commercial activities that seemingly would have occurred without a university's patentable contributions. Indeed, some universities have gone further, on occasion selling their rights to patent trolls who make their money by threatening to disrupt or prevent commercialization. Senators Bayh and Dole would, one suspects, not be pleased by some university actions their law has enabled.

Eyewitness testimony, the subject of Judge Jed Rakoff and Elizabeth Loftus's essay, is the single most common factor contributing to wrongful convictions for serious crimes. Rakoff and Loftus briefly discuss why eyewitness testimony is such powerful evidence before reviewing what we know about the causes of mistaken eyewitness identifications. They then explore efforts that have been made to increase eyewitness accuracy and to help factfinders assess the strengths and weaknesses of eyewitness testimony in trials. Their essay not only reports ways in which the social sciences have been used to identify weaknesses in eyewitness testimony and ways to ameliorate them, but also documents ways in which this knowledge has led to procedural reforms designed to increase the accuracy of eyewitness testimony and the ability of jurors to evaluate it.

A key distinction made by the authors is the difference between system variables and estimator (or witness) variables. The former has to do with the way eyewitness identifications are elicited: how lineups are constructed, for example. Problems of this sort are relatively tractable, and in many states, scientific findings have led to promising procedural change. Problems posed by the latter – that is, by weaknesses inherent in human observation and memory – pose far more difficult challenges. The best we may be able to do, Rakoff and Loftus suggest, is to educate judges and jurors on factors that, if present, make eyewitness identifications problematic so that they can do a better job of weighing an identification's probative value.

Jennifer Mnookin succeeds in presenting, in remarkably brief compass, an informative account of the state of forensic science today. After effectively acquainting readers with the forensic identification sciences, she highlights issues that are now dominating discussions both within the forensic science community and among the leading critics of forensic science procedures, protocols, and modes of testifying. Mnookin herself has been an important and respected participant in these discussions, especially as they relate to friction ridge (fingerprint) identifications, and one can see why. Her positions are not dogmatic, nor are they entirely critical; rather they both recognize deficiencies in forensic science technologies and ways of testifying, and acknowledge efforts being made, including efforts by forensic science practitioners, to improve the quality and characterizations of the forensic science evidence they offer.

She supports her claim that one may see the current state of the forensic identification sciences as a glass half empty or half full by reference to a pair of contrasting bite mark identification cases that arose in the states of Connecticut and Pennsylvania within months of each other. In the Connecticut case – a review of a 1991 murder conviction in which bite mark evidence played a major role – the defense, the prosecution, and the scientist who presented the original bite mark evidence agreed that the bite mark identification was worthless, with the expert even calling his earlier testimony "junk science." Combined with corroborating DNA evidence, the judge vacated the murder conviction and reopened the case. In the Pennsylvania case, the trial judge refused to even hold a full hearing to determine if the bite mark evidence offered by

the prosecution was sufficiently reliable to be admitted, citing precedent that allowed it. The two cases may be distinguished, but the weaknesses of bite mark evidence are so well known that if it is regarded as sufficiently reliable to be admitted, judicial barriers against other frequently offered forensic science evidence would seem unlikely, no matter how frail the evidence's scientific underpinnings. Mnookin believes, however, that further reform is possible, and identifies collaboration between research scientists and stakeholders in the legal system as the best hope for transformative change.

Because uncertainty attaches to all forensic science claims, effectively communicating levels of certainty to factfinders is crucial to accurate fact-finding. Joseph Kadane and Jonathan Koehler present results from an experiment that tests whether the words that fingerprint examiners use to express their conclusions affect the weight that laypersons give reports of possible matches. They find that the two most scientifically defensible ways of reporting on fingerprint comparisons, neither of which claims that two fingerprints indisputably match, have the effect of moderating judgments, when compared to other ways that examiners might express opinions that two fingerprints match. If an examiner is willing to say that she thinks two fingerprints match, respondents are not sensitive to differences in the language used to fortify that opinion.

This study is important early research, an original study using a brief written transcript and nondeliberating mock jurors, but it is a first step. Research in other areas where social science findings have affected legal procedures, such as the eyewitness reforms discussed in the Rakoff-Loftus essay, began with similar small steps, followed by more elaborate studies in the laboratory and in the field. Kadane and Koehler's findings are intriguing enough that they should stimulate research to confirm what they have found, helping both scientists and the legal system to hone in on ways that protocols for communication can improve practice.

Nancy Gertner and Joseph Sanders begin their essay by suggesting that two principal goals of judicial trials, accuracy and fairness, are not consistent. Accuracy references an objective standard, while fairness lies in the eyes of the beholder. Gertner and Sanders cite research suggesting that, consistent with the American model of adversary litigation, people see decisions that affect them as fairer when they have had an opportunity to provide information to the decision maker and to have their stories heard. Accuracy, on the other hand, is thought by some as likely to increase when an expert judge closely controls proceedings and witnesses are not identified with parties. When scientific matters are at issue, not only does party control lead to the biased selection of experts who may not be representative of the best available expert opinion, but serving as a party witness can color expert evaluations and the way experts report their findings, even when they think they are being objective.

Having laid out the potential tension between accuracy and fairness and the research pointing to it, Gertner and Sanders explore suggested reforms aimed at enabling more accurate evaluations of scientific evidence within the general confines of the American adversary system. These include readjusting the order of testimony so that opposing experts testify in temporal proximity to each other; adopting the Australian procedure of "hot tubbing," in which experts appear together before the factfinder to present and discuss their differing views; and making changes in jury procedure likely to increase the ability of jurors to understand expert testimony and better judge where the weight of the scientific evidence lies. The authors explore not

just the potential benefits from such changes but also potential downsides and difficulties of implementation. Implicit in the Gertner-Sanders essay is a message more explicitly stated in other contributions: while we can be confident that some reforms, mainly relating to jury management, are likely to improve the evaluation of expert testimony, we need more research that targets other reforms, particularly those relating to expert selection, information sharing, and the presentation of expert testimony.

Daniel Rubinfeld and Joe Cecil discuss the core challenge that scientific evidence often poses for judges and juries: namely, difficulties in understanding which side to believe when the parties' experts present conflicting scientific testimony and the triers, unschooled in the science, have in their prior knowledge little basis for preferring one side's analysis to the other's. The authors review three methods the law has developed to help courts better evaluate science: court appointed experts, court appointed advisors, and special masters. Court appointed experts, like the parties' experts, evaluate the relevant evidence and may testify in court, subject to cross-examination. Their apparent neutrality is thought to make their views particularly influential if they testify, which in turn means that their findings may stimulate settlements rather than be a precursor to testimony. Court appointed experts may also contribute without rendering opinions by, for example, getting the parties to agree on a common data set or on the methods to be used in their analyses. Court appointed science advisors serve a function much like a judge's law clerks, except they assist the judge in evaluating the scientific evidence in the case while the ordinary law clerk assists by assembling relevant legal materials and aiding in opinion writing. Special masters fill a judge-like role. They can hear evidence, sort through material, help with discovery, and issue recommended findings for a judge to consider. Where a case turns on scientific evidence, they can be chosen for their expertise in the relevant science.

None of these procedures is in common use, and although they are attractive options, they also have, as Rubinfeld and Cecil point out, potential shortcomings. These include the extra costs they impose on parties and the possibility that they may have undue influence on final results, particularly if the science is not settled. Experts may be unbiased in their relationship to the parties, but they may favor or deplore particular scientific methods or schools of thought.

Valerie Hans and Michael Saks begin their essay by noting the fundamental paradox that motivates several of the essays: "those with the power and duty to evaluate expert testimony possess less knowledge of the specialized subject matter at issue than that possessed by the experts whose testimony they are evaluating." Moreover, "Expert evidence must be prescreened for nonexpert jurors by nonexpert judges." If this is not trouble enough for the legal system, Hans and Saks point to general shortcomings of human reasoning, including the degree to which rationality may be subverted by biases relating to how information is acquired and the use of heuristics. Yet the Hans and Saks essay is more optimistic than pessimistic about the capacity of judges and juries to deal with expert scientific evidence. They point to the importance of factfinder neutrality in evaluating conflicting expert claims and to the ways in which the organization of trials and collective decision-making work to foster careful processing of information.

Perhaps most striking in the Hans and Saks essay is the number of studies they can reference that provide an empirical basis for procedures and reforms that are likely to enhance the capacity of jurors and judges to understand and rationally eval-

uate the claims experts make. Also striking is how few of the studies have been replicated to create a robust body of research, allowing an observer to say with confidence, "this will work" rather than "this appears promising." Their conclusion, thus, is hard to dispute: "We must collect data and run experiments; that is, we should take a scientific approach to deciding on those reforms that will best enable judges and juries to cope with modern scientific evidence."

In their closing essay, David Baltimore, Judge David S. Tatel, and Anne-Marie Mazza highlight the challenges posed by the distinct cultures of science and the law and discuss one of the most important recent developments in efforts to bridge gaps between these cultures: the creation of new, broadly representative institutions that bring members of both cultures together to work cooperatively on issues that are raised at their intersection. Baltimore and Judge Tatel currently cochair one of the most important manifestations of this effort: the Committee on Science, Technology, and Law (CSTL), a new standing committee that serves under the auspices of the National Academies of Sciences, Engineering, and Medicine. In their essay, Baltimore, Tatel, and Mazza describe the concerns that inspired the creation of the CSTL and the legal backdrop that helped stoke these concerns. They then highlight some of the CSTL's accomplishments, including its influence on rule-making and public policy and the establishment, under its auspices, of a committee that took a hard look at the scientific foundations of the different forensic sciences, an effort yielding a critical report that sparked an ongoing national conversation about the forensic sciences, affecting both the legal and scientific communities. Other efforts have been similarly well received. Together with ongoing research, bringing experts of this sort together has an important role to play in improving the quality of the science offered to courts and the ability of courts to intelligently evaluate that science.

As editors of this volume, we are delighted by the range of new and thoughtful insights about the relationship between science and the legal system represented by the essays in this collection. The authors do not provide solutions to all of the challenges presented by the interface between science and the legal system. The gaps, pushbacks, and procedural obstacles will continue to require attention, borrowing from Mnookin's characterization, to fill the science-law glass. They do, however, provide reasons for optimism about future collaboration between science and law.

AUTHOR BIOGRAPHIES

SHARI SEIDMAN DIAMOND, a Fellow of the American Academy since 2012, is the Howard J. Trienens Professor of Law and Professor of Psychology at the Northwestern University Pritzker School of Law, and Research Professor at the American Bar Foundation. She is the author of the *Reference Guide on Survey Research* in the *Reference Manual on Scientific Evidence* (3rd ed., 2011) and *The Multiple Dimensions of Trial by Jury: Studies of Jury Behavior* (2016, in Spanish) and editor of *Trademark and Deceptive Advertising Surveys: Law, Science, and Design* (with Jerre B. Swann, 2012).

RICHARD O. LEMPERT, a Fellow of the American Academy since 1993, is the Eric Stein Distinguished University Professor of Law and Sociology, emeritus, at the University of Michigan. He is the author of *A Modern Approach to Evidence: Text, Problems, Transcripts, and Cases* (with Samuel Gross et al., 5th ed., 2013) and *An Invitation to Law and Social Science* (with Joe Sanders, 1986) and editor of *Evidence Stories* (2006).

[1] Tal Golan, *Laws of Men and Laws of Nature: The History of Scientific Expert Testimony in England and America* (Cambridge, Mass.: Harvard University Press, 2004), 20.

[2] *Alsop v. Bowtrell* (1619) Cro. Jac. 541; 79 ER.

[3] Carol A. G. Jones, *Expert Witnesses: Science, Medicine, and the Practice of Law* (Oxford: Oxford University Press, 1994), 20.

[4] Golan, *Laws of Men and Laws of Nature* [see note 1], 49.

Science, Common Sense & Judicial Power in U.S. Courts

Sheila Jasanoff

Abstract: Courts routinely resolve factual disputes as an adjunct to settling legal controversies, and such fact-finding frequently involves scientific and technical evidence. It is important to ask what intellectual resources judges bring to this task. Instead of assessing how much science judges know or understand, this essay focuses on the judge's role in articulating and reinforcing prevailing cultural attitudes toward science. Background judicial assumptions matter at three significant junctures. First, judges maintain the lay-expert boundary by deciding whether an issue demands expert testimony at all. Second, judges act as epistemological gatekeepers, by determining which expert claims and ways of reasoning are entitled to deference and which are not. Third, judges decide how to classify and categorize things of uncertain ontological status as a prelude to applying legal rules. Each kind of decision offers a window into judicial common sense, a relatively neglected topic in studies of law and science.

SHEILA JASANOFF is the Pforzheimer Professor of Science and Technology Studies at Harvard University. She is the author of *The Ethics of Invention* (2016), *Science and Public Reason* (2012), and *Designs on Nature: Science and Democracy in Europe and the United States* (2005).

The courtroom is a space of reenactment. Something happened in the world to awaken society's demand for moral reckoning: someone must be blamed, someone punished, someone rewarded for exceptional enterprise, someone, if possible, made whole. Whether the event was a deadly assault or the misappropriation of private funds through an elaborate Ponzi scheme or a scientific discovery giving rise to intellectual property claims, the legal process offers an opportunity to replay the sequence of events before an authority capable of making binding judgments that satisfy our collective sense of order, compassion, or moral indignation. Such weighty decisions demand a full-blown commitment to factual truth, for without a baseline of agreed upon facts, no judgment could satisfy the world's demands for justice.

Courts can be seen in this sense as sites of translation. What happened back there and then must be replayed as accurately as possible here and now before an empowered moral adjudicator, a judge,

doi:10.1162/DAED_a_00517

usually supported in U.S. lower courts by a fact-finding jury. Like a pointillist painting decomposed into its individual dots and pixels of paint, each moment, each unit of action from the bygone event must be brought into the adjudicatory setting, physically or verbally, in a form sufficiently reliable to render moral evaluation both possible and plausible. Each element, then, must be transported before the eyes of the adjudicator in trustworthy form, a form recognizably related to the reality of the circumstances in question. No wonder, then, that a murder trial can consume months of preparatory time, a corporate financial scandal can take years to unravel, and a regulatory or patent controversy can take seven years or more to journey to the Supreme Court. No wonder, too, that the rules of translation by which the external drama is brought in and reenacted in contexts of adjudication have attracted so much attention from legal analysts.

Scientific evidence presents special problems of translation. First, science itself is already a form of translation: it is a means of making the facts of nature knowable in human terms, through instrumental measurements, visual or quantitative representations, and specialist discourses that enable followers to build on findings that have gone before. Second, when serving the purposes of the law, science and its associated technologies offer an especially powerful means for bridging time and space, as warranted truth-telling mechanisms that can, when properly used and interpreted, bypass distortions produced by human memory or motives. Yet science cannot speak for itself to a legal factfinder. Science's gaze on matters in dispute is always at a remove, transmitted through intermediaries, both human and nonhuman, that stand in for what actually *is*. When scientific evidence is introduced in court, there is thus a double challenge: the presentation must close the gap between the original action and its courtroom replication (for example, by establishing a chain of custody for physical samples) and it must persuade the court that science's findings relate truthfully and reliably to the events, actions, intentions, and consequences that are the subject matter of adjudication.

The primary social innovation through which the law has sought to accommodate science is the figure of the expert witness. Rule 702 of the Federal Rules of Evidence provides that a person qualified by "knowledge, skill, experience, training, or education" can offer specialized testimony to facilitate a court's determination of scientific or technical facts. The expert testifies to the authenticity and meaning of the traces left by the questioned actions, thereby bridging the gap between the unrecorded past and its present reenactment. This performance entails a second-order problem that has preoccupied the law for more than two hundred years.[1] How can the legal factfinder be sure that the expert is offering dependable testimony and not unsubstantiated personal opinion or, worse, false, fraudulent, or misleading views clothed in the authority of expertise?

In this essay, I focus not on the reliability of expertise, but on the judge's role in articulating and reinforcing prevailing cultural attitudes toward science. This topic has received relatively little attention from legal practitioners and scholarly commentators. Yet judicial thinking is of paramount importance in three ways. First, judges consider and ratify how scientific and legal authority should work vis-à-vis each other, for instance by determining whether an issue does or does not demand expert testimony. Second, judges play the part of epistemological gatekeepers. The judge's eye determines which expert claims are entitled to consideration in the courtroom, or not, thereby privileging certain ways of knowing above others. Third, and perhaps least visibly, judges exercise ontolog-

ical power by deciding how to classify and categorize things for purposes of legal decision-making.

In making all three sets of moves, courts operate to some extent as amplifiers of common sense, importing widely held cultural ideas about how things work into their assessments of both the necessity for and the reliability of scientific and technical expertise. Though tacit and informal, such judgments are neither wholly subjective nor arbitrary. They are rooted in engrained collective beliefs, a common sense that has power precisely because it operates below the level of conscious argument, in a register of cultural familiarity, and hence is not open to questioning, indeed is accepted as integral to law.

In an influential essay, the anthropologist Clifford Geertz urged his fellow cultural analysts to view common sense as an ordered system of thought, on a par with more formal systems such as "physics, or Islam, or law, or music, or socialism."[2] Common sense, in Geertz's telling, fills in the gaps of experience, when conventional explanations and classifications fail, and it does so in ways that are culturally intelligible, widely shared, and hence unquestioned by members of a given society. Boundary-crossing anomalies, Geertz suggested, are treated differently in different cultures. Intersexuality, to take one example, is known in all human societies, but it is variously classified as horror, wonder, or simple biological error because different shared assumptions about the nature of sexuality condition responses to the apparent anomaly of not being either simply male or simply female. Geertz concluded that, "Common sense is not what the mind cleared of cant spontaneously apprehends; it is what the mind filled with presuppositions . . . concludes."[3] Through an analysis of significant Supreme Court decisions, this essay probes the presuppositions about science and technology, and their uses as evidence, that fill the minds of the federal judiciary.

What qualifies an expert's testimony as good enough to count as pertinent evidence? The U.S. Supreme Court wrestled with this question in three landmark evidence decisions of the 1990s, beginning in 1993 with *Daubert v. Merrell Dow Pharmaceuticals, Inc.*[4] In that first and still most significant decision, the Court held that the earlier rule for the admissibility of scientific evidence in federal proceedings, derived from a 1923 appellate decision in a murder trial, *Frye v. United States*, had been superseded by the Federal Rules of Evidence.[5] The *Frye* standard turned on whether a novel scientific procedure enjoyed general acceptance in the relevant scientific community. The Federal Rules of Evidence, as interpreted in *Daubert*, did not endorse this one factor test.[6] More pointedly, the Court reminded judges that they were responsible for acting as gatekeepers with respect to proffered expert testimony and offered guidance on what that meant. Judges should think like scientists in assessing the relevance and reliability of scientific evidence, using the same criteria that scientists would apply. While cautioning against treating them as a "checklist," the Court named four criteria that instantly became, to some degree, canonical: is the claim falsifiable and has it been tested; was it peer reviewed; has an error rate been determined; and has the underlying science won general acceptance?

Following *Daubert*, the judge's understanding of what science is, how it works, and what constitutes legitimate expert representations of scientific knowledge became a decisive influence on determinations of admissibility. What, though, did this shift mean in terms of "law's knowledge"?[7] Did science's ways of knowing indeed displace traditional modes of judicial reasoning, or was some more complex al-

chemy at work in the translation exercise that *Daubert* so radically reconfigured? Did particular traits of judicial epistemology, particular styles of reasoning, or ways of assessing the facts of the world gain power and influence in the post-*Daubert* adjudicatory environment? *Kumho Tire v. Carmichael*, the last of the *Daubert* trilogy, offers particular illumination.[8]

On July 6, 1993, Patrick Carmichael, one of the plaintiffs in *Kumho*, was driving a minivan when the right rear tire blew out, killing one passenger and severely injuring several others. The plaintiffs claimed that the blowout was due to a defect in the design or manufacture of the failed steel-belted radial tire. Their case rested to a significant degree on the testimony of Dennis Carlson Jr., a mechanical engineer and professed expert in tire failure analysis, who offered his informed opinion that the blowout was not caused by ordinary wear or misuse, but rather by a design defect.

Through visual and tactile inspection, Carlson concluded that a manufacturing defect had caused the tread to separate from the body, or "carcass," of the tire, despite evidence that the tire was seriously worn and had been inadequately repaired for punctures on two occasions.[9] The district court mechanically applied the four *Daubert* criteria to Carlson's evidence and found it inadmissible. The Eleventh Circuit Court of Appeals reversed on the ground that the *Daubert* standard applied only to scientific, not technical, evidence, and the Supreme Court, under Chief Justice Rehnquist, agreed to review that decision. The questions before the Court were whether *Daubert*'s gatekeeping criteria applied only to scientific evidence or also to technical and other nonscientific expert evidence; and, if so, whether the four *Daubert* criteria could be used to assess reliability in this case. The Court ruled positively on both counts, reversing the Eleventh Circuit ruling.

In the original trial and first appeal, Dennis Carlson's legitimacy as an expert had not been in question. But it was not obvious to the courts what kind of expert he was and, consequently, whether his kind of knowledge could be held to the *Daubert* standard for scientific expertise. Judge Stanley Birch, writing for the Eleventh Circuit, ruled that this determination was crucial. "In short," Birch concluded, "a scientific expert is an expert who relies on the application of scientific principles, rather than on skill- or experience-based observation, for the basis of his opinion."[10] Citing a Sixth Circuit decision to support this distinction, Birch revisited that court's analogy, in which a hypothetical jury needs an explanation of a bumblebee's ability to fly.[11] You might bring in an aeronautical engineer, the Sixth Circuit mused, to explain general principles of flight that could be applied to the bee. Even if such an expert had never seen a bumblebee, the testimony could still be admitted as relevant evidence. On the other hand, the testimony of a beekeeper with no scientific training could also plausibly tell the jury, on the basis of firsthand observations, that bumblebees always take off into the wind. "In other words," the Sixth Circuit concluded, "the beekeeper does not know any more about flight principles than the jurors, but he has seen a lot more bumblebees than they have."[12] Here, the beekeeper's experience is seen as different in degree, but not in kind, from that of a juror, and is entitled to be heard for that very reason: the beekeeper knows relevant facts better than any juror. This is not so for the aeronautical engineer, who knows nothing about bees in particular and hence must draw on certified theoretical knowledge for authority.

Carlson, by this reckoning, presented a conundrum. With formal degrees in mechanical engineering *and* ten years of experience in tire testing at Michelin, Carlson offered testimony that was hard to classify in terms of the beekeeping analogy.

Judge Birch wondered, "is the testimony at issue in this case more like that of a beekeeper applying his experience with bees or that of an aeronautical engineer applying his more generalized knowledge of the scientific principles of flight?"[13] Despite Carlson's engineering qualifications, Birch concluded that he was, in terms of the issue at hand, a beekeeper of tire failures: "Like a beekeeper who claims to have learned through years of observation that his charges always take flight into the wind, Carlson maintains that his experiences in analyzing tires have taught him what 'bead grooves' and 'sidewall deterioration' indicated as to the cause of the tire's failure."[14] Ergo, Birch reasoned, Carlson's testimony fell outside of *Daubert*'s scope – in the realm of experience rather than science – and the district court therefore erred in applying the *Daubert* criteria and ruling his evidence inadmissible.

The Eleventh Circuit's attempt to draw a bright line between science and nonscience flies in the face of much historical work in science and technology studies showing that, in the conduct of science, there is no essential distinction between theory and practice, or "head" and "hand" in the terminology of historian Steven Shapin and sociologist Barry Barnes.[15] Such demarcations are culturally produced and pedagogically transmitted rather than intrinsic to the scientific enterprise. The Supreme Court did not cite such insights, but came to similar conclusions from different assumptions about how to articulate a sensible demarcation between science and nonscience.

During oral argument, Chief Justice William H. Rehnquist signaled his discomfort with any categorical distinction between science and expertise. "All right," he summed up with more conviction than elegance, "and then you'd also agree that there isn't a rigid categorization as between science or not where you could say the *Daubert* test is or is not useful. The answer is both within and outside something that the Harvard University would call science or something. I mean, sometimes within that, sometimes outside of it . . . *Daubert*'s helpful, sometimes it's not helpful."[16]

Crucially, Rehnquist indicated that it is the judge who decides on a case-by-case basis when *Daubert* is "helpful" and when it is not – not "the Harvard University" nor the academic scientific establishment. This point was brought home by Judge Jed S. Rakoff during the discussions leading to this issue of *Dædalus*. He noted that judges routinely make distinctions among *Daubert*'s four criteria based on their preconceived understandings of what is or is not germane to doing good science:

> I think this error rate one is often not considered a requirement. There are many kinds of science that – they don't have a known error rate, and I think *many judges will accept* that that's not dispositive. . . . But with respect to whether it's been tested or not, most judges seem to believe that, "God, if it isn't – hasn't been tested, how could it be called science?" So, yeah, that one is taken as a sine qua non. Has it been peer reviewed and the subject of publication? Well, if it hasn't been that, then it's just someone's . . . idea – that we have no idea whether it's ever been put to the test, and *the test there is very similar to the legal tests of cross examination*. So *it comes naturally to judges* to say, "If it hasn't been peer reviewed and publicized, that's . . . pretty damning." The error rate, different – I don't think more judges regard that as a sine qua non, *and then the fourth is, of course, the old-fashioned Frye test*, whether it's generally accepted, and the question, always, there is what's the relevant group.[17]

The passage as a whole illustrates the commonsensical mindset with which judges decide how to apply *Daubert*, a process that foregrounds longstanding judicial intuitions about what makes any claim stronger or weaker than another. Particu-

larly noteworthy in this text is the equating of peer review with cross-examination, a method of adversarial questioning deeply familiar to judges and one long seen as capable of separating the wheat of truth from the chaff of false pretenses.

Later in the *Kumho* oral argument, Rehnquist clarified his position regarding expert evidence: namely, that inductive arguments are insufficient unless they are, in effect, theory-laden.

> [I]n my mind, anyway, I think the hardest question for you would be, you'd say, well, look, there is a theory going on here that in the absence of these four specific factors, not any kind of abuse but four kinds, beading, flange, whitewall discoloration, and some other thing, that your expert seems to say, in the absence of those four things, it must have been defect.
>
> And immediately *a common sense person thinks,* what? You mean nails couldn't be an abuse? You mean, it's bald couldn't be an abuse?
>
> And the expert says – if the expert then says, well, I have a lot of experience at this, you say, wait a minute. You couldn't have seen hundreds or thousands of tires that have had two nails – you know, two nails driven into them, and they're bald, and they've gone 100,000 …that's impossible.
>
> You're going on some theory, and if you're going on some theory, you tell me who else believes that theory.[18]

Implicit in Justice Rehnquist's thinking, as in Judge Birch's, is the idea of the putative "common sense person" as an expert on things-in-the-world, and a person whom the judge is entitled to represent when elucidating such everyday understandings. In his spontaneous dramatization of expertise encountering lay skepticism, the Chief Justice in effect tests the limits of the expert's reasoning, as well as the improbable certainty of his experience-based claims, by constructing alternative, common-sense scenarios that display the gaps between Carlson's observations and the conclusion drawn from them. To support a claim on the basis of experiential knowledge, Rehnquist's imagined interlocutor insists, the expert must be "going on some theory," because only such a theory could rule out all other intervening causal stories (such as the nails or the baldness); and then the expert had better be able to marshal the resources of a like-minded community ("you tell me who else believes that theory"). If such support is not forthcoming, Rehnquist implies, then that expert's gaze is no more reliable than anyone else's.

Behind Rehnquist's questioning is classic Humean skepticism, an assumption that a finite number of observations of other tires could not possibly provide a firm basis for conclusions regarding the one that failed. The only legitimate foundation for so particular a claim must be a general theory, and here Rehnquist reverts back to the familiar comfort of the *Frye* rule. If there is an applicable theory, then others should also believe in it; in other words, it should be generally accepted.

In deciding *Kumho,* the Court unanimously agreed that no *a priori* boundary between science and engineering or other forms of expertise was practically workable: "Finally, it would prove difficult, if not impossible, for judges to administer evidentiary rules under which a gatekeeping obligation depended upon a distinction between 'scientific' knowledge and 'technical' or 'other specialized' knowledge. There is no clear line that divides the one from the others."[19] Illustrating a judicial predilection for citing legal authority even for matters of epistemic principle, the Court turned to the great common law jurist Learned Hand for the proposition that experts may come to their conclusions through the use of "general truths derived from…specialized experience."[20]

But it was in part three of the opinion, authored by Justice Stephen Breyer, that the majority most clearly articulated its epistemological sensibilities. Ostensibly instructing the trial court on how it could reasonably have applied the *Daubert* criteria to Carlson's testimony, Justice Breyer never mentioned the four tests. He instead conducted, in effect, his own virtual inspection of the contested tire; significantly, the opinion even included a picture from a manual on how to buy and care for tires. The conclusions reached by the tire expert's eye fell short in the light of the judge's (presumably more rigorous) reexamination of the evidence:

> The [trial] court could reasonably have wondered about the reliability of a method of visual and tactile inspection sufficiently precise to ascertain with some certainty the abuse-related significance of minute shoulder/center relative tread wear differences, but insufficiently precise to tell "with any certainty" from the tread wear whether a tire had traveled less than 10,000 or more than 50,000 miles.[21]

We see here the law's age-old reliance on direct eye-witnessing as the means through which events are most reliably reconstructed in the courtroom – but with a twist.[22] Carlson's spurious precision failed to meet the common-sense standard of "intellectual rigor" that Justice Breyer and his coauthors deemed necessary to rule out alternative causes.

The *Daubert* trilogy tilted epistemic authority subtly but surely in favor of how judges see and know the world, including how they imagine science itself, when they are prepared to substitute their own authority for that of an expert witness, and how they classify the products of science and technology. These judgments are pervasive, cutting across many domains of law that are not normally seen as ripe for epistemic analysis; for example, environmental law, intellectual property law, and constitutional law. Yet in high-profile cases in all these areas, the ultimate legal judgment has turned on how the courts, including especially the Supreme Court, analyze the things that science and technology introduce into the world. Once again, these are decisions in which judicial common sense governs, though the foundations of such intuitions are seldom questioned or laid bare for critical inquiry. Examples from recent case law illustrate these points.

Environmental law. Few areas of modern law rely as much on the scientific assessment of causes as environmental regulation and the repeated challenges against it. Causes and consequences are difficult to establish with any certainty. It is clear from the long record of environmental litigation that repose on technical issues ultimately results less from agreements about what is true than from parties' acceptance that scientific assessment procedures were properly followed, including those for soliciting expert advice and subjecting it to the scrutiny required by applicable statutory mandates.

Environmental law runs into special difficulties when regulatory action is directed toward previously unrecognized hazards. In these cases, the regulator often confronts an entity or agent that was either not known at all (such as small particulate matter deemed since the late 1990s to be substantially responsible for urban respiratory disease), or is shown to have unsuspected properties that make it no longer suitable for its original purposes (for example, lead as antiknock agent, DDT as insecticide, thalidomide as anti–morning sickness drug, or atrazine as weed killer). In such cases, questions about the science become interlaced with politics. Huge stakes may hang on whether a product crosses the line from safe to dangerous or, indeed, is recognized at all as a potential regulatory target.

The long-running U.S. debate on climate change illustrates how environmental science is vulnerable to concerted attack when new, scientifically certified objects and phenomena threaten settled lifestyles. The first two decades of the twenty-first century saw repeated reversals in federal policy based on the political alliances of the administration in power, particularly along the dividing line between fossil fuels and renewable energy. For the most part, these conflicts played out at the level of science and regulatory policy at the Environmental Protection Agency (EPA), but they spilled into courts in one landmark case, *Massachusetts v. EPA*, which also serves as a kind of instruction manual on how judges negotiate the competing claims of science and law in rendering the facts of nature tractable for moral adjudication.[23]

In this case, the majority deferred to science, as the EPA also had, in accepting "the existence of a causal connection between man-made greenhouse gas emissions and global warming." But unlike the EPA, the Court also concluded that the language of the Clean Air Act was expansive enough to admit new entities like greenhouse gases into the definition of "air pollution": "While the Congresses that drafted §202(a)(1) might not have appreciated the possibility that burning fossil fuels could lead to global warming, they did understand that without regulatory flexibility, changing circumstances and scientific developments would soon render the Clean Air Act obsolete. The broad language of §202(a)(1) reflects an intentional effort to confer the flexibility necessary to forestall such obsolescence."[24] Resolving the definitional question also resolved the issue of the EPA's authority to act: "Because greenhouse gases fit well within the Clean Air Act's capacious definition of 'air pollutant,' we hold that EPA has the statutory authority to regulate the emission of such gases from new motor vehicles."

Justice Antonin Scalia, in a sharply worded dissent, disagreed with the majority's construction of the act and urged a more prosaic reading of the term "air pollutant." He found less certainty in the science than his colleagues did, but just as importantly, he concluded that the EPA had rightfully interpreted the words of the Clean Air Act as not requiring the regulation of greenhouse gases. Scalia's turn to common sense took the form of insisting that the language of the law be given its plain meaning:

> We need look no further than the dictionary for confirmation that this interpretation of "air pollution" is eminently reasonable. The definition of "pollute," of course, is "[t]o make or render impure or unclean." Webster's New International Dictionary 1910 (2d ed. 1949). And the first three definitions of "air" are as follows: (1) "[t]he invisible, odorless, and tasteless mixture of gases which surrounds the earth"; (2) "[t]he body of the earth's atmosphere; esp., the part of it near the earth, as distinguished from the upper rarefied part"; (3) "[a] portion of air or of the air considered with respect to physical characteristics or as affecting the senses." *Id.*, at 54. EPA's conception of "air pollution" – focusing on impurities in the "ambient air" "at ground level or near the surface of the earth" – is perfectly consistent with *the natural meaning of that term*.[25]

Faced with the ontological problem of slotting a new physical entity – "greenhouse gases" – into a preexisting statutory framework, the justices divided in their conclusions, but each position rested on the author's own tacit sense of how the law-science relationship should properly work. For the majority, it made sense that science declares the state of how things are, and it is only natural to interpret broad legal language to accommodate changes in our understanding of the world. For Justice Scalia, a strong advocate for the sovereignty of the legal text, it was just as nat-

ural (or commonsensical) to insist that words, first of all, be given their ordinary meaning.[26] If those "natural" meanings reasonably supported the agency's decision not to recognize a new regulatory object, then no amount of scientific urgency could undermine that judgment. The remedy, if any, would have to come from the legislature that wrote the law, the only body entitled to change the words to permit a new reading.

Intellectual property law. Ontological judgments are the basic stuff of intellectual property decisions, since at the core of most awards or denials of such rights are determinations whether something new (or, in the case of copyright, original) has been created and, if so, whether it is the kind of thing for which the award of such rights was meant. In the case of patents, both judgments reveal tacit judicial understandings of what inventiveness means and where the boundary lies between nature and human artifice, along with beliefs about the right relationship between scientific and legal innovation.

Thus, in *Diamond v. Chakrabarty*, the landmark 1980 decision in which a divided Supreme Court held that human-made living organisms are no different from nonliving ones for purposes of patenting, Chief Justice Warren Burger's opinion cast the law's role as essentially passive.[27] Like the majority opinion in *Massachusetts v. EPA* almost thirty years later, *Chakrabarty* construed the governing law as expansive enough to accommodate changes in science. Congress, the Court famously held, "plainly contemplated that the patent laws would be given wide scope," so that patents could be granted for "anything under the sun that is made by man." At the same time, the Court positioned itself as powerless to change the course of scientific or technological progress: "legislative or judicial fiat as to patentability will not deter the scientific mind from probing into the unknown any more than Canute could command the tides."[28] This was a remarkable bit of rhetorical jiujitsu in a decision widely regarded as having enabled the modern biotechnology industry to come into being, and it was justified in part by invoking a trope of demonstration through ordinary empirical witnessing: the king at the shore powerless to hold back the sea from advancing.

The importance to courts of the notion of plain, unobstructed seeing shines through in another patent decision overturning years of settled legal practice: the Supreme Court's 2013 decision in the Myriad Genetics case, ruling that human genes are not patentable.[29] Here, in a case challenging patents that Myriad held on human breast cancer genes, the Justice Department and the American Civil Liberties Union (ACLU) presented the Court with metaphors that would make plain why only one conclusion was reasonable. The genes that Myriad had isolated, petitioners claimed, could be seen by anyone who cared to look; it took no special inventiveness to discern them. To make this argument stick, the Justice Department invented a hypothetical instrument – the "magic microscope" – arguing: "[I]f an imaginary microscope could focus in on the claimed DNA molecule as it exists in the human body, the claim covers ineligible subject matter."[30] Chris Hansen, lead lawyer for the ACLU, opted in oral argument for a still more elemental metaphor: gold, with its connotations of extraction and mining. Finding a method of extracting gold, Hansen said, might entitle one to a patent, as would finding a new use, such as "a new way of using gold to make earrings."[31] But the gold itself would not be patentable and neither are genes extracted from the human body.

Unlike the reference to King Canute in *Chakrabarty*, which echoed an *amicus* brief by the biotechnology company Genentech, neither the magic microscope nor the gold

analogy survived into the Court's gene patenting decision. The moves that ACLU attorney Hansen made to classify genes as products of nature did, however, resonate. With the same matter-of-factness conveyed in the ACLU's oral argument, the Court ruled that "Myriad did not create anything. To be sure, it found an important and useful gene, but separating that gene from its surrounding genetic material is not an act of invention."[32] If nature was the initial inventor, then no amount of brilliance, effort, or innovation could render nature's work patentable. Put differently, the Court concluded: "discovery, by itself, does not render the BRCA genes 'new…composition[s] of matter,' §101, that are patent eligible."[33] And the key to distinguishing between invention and discovery remained the act of seeing: anyone, after all, could see that the "location and order of the nucleotides [in an isolated gene] existed in nature before Myriad found them."[34] By contrast, synthetic complementary DNA (cDNA) could be patented because it is made up of a nucleotide sequence that does not visibly exist within the body.

Constitutional law. In an era in which human lives are ever more intimately entwined with the products of science and technology, ontological judgments have begun to figure with increasing frequency in constitutional decision-making. Back in 1967, in what now feels almost like ancient history, the Supreme Court decided 7 to 1 in *Katz v. United States* that a warrantless wiretap violates the Fourth Amendment.[35] A physical intrusion was not deemed necessary for constitutional purposes; it was sufficient that the defendant had sought to reserve the space as private. It was in this respect, Justice John Harlan concurred, an area where, as in a home but not in a field, "a person has a constitutionally protected reasonable expectation of privacy."[36] The telephone booth was transformed, in the eye of the Court, into an enclosed space, similar to a room, whose walls should have provided safeguards against the intrusive, if metaphorical, "presence" of the wiretapping machine.

Developments in many areas of engineering and technology (such as nanotechnology, gene editing, robotics, and artificial intelligence) are further blurring boundaries between taken-for-granted classifications that once provided clear baselines for constitutional jurisprudence. At stake are questions about the division between nature and artifice, life and death, and human and nonhuman. Is a cell line sufficiently continuous with the human body it came from to deserve some degree of special treatment, such as informed consent to being used in research?[37] What sorts of personal rights extend to "data subjects," for example, the right to be forgotten?[38] What would it mean for robots to be classified as "electronic persons," with explicit rights and obligations? Questions such as these are bound to proliferate in coming decades, focusing renewed attention on the intellectual resources with which courts approach these novel tasks of boundary drawing.

Such issues are already being addressed by U.S. high courts. An instructive example is the Supreme Court's 2014 decision in *Riley v. California*, holding that the Fourth Amendment protects against warrantless searches of cell phones.[39] While this decision can be seen as a principled extension of earlier decisions such as *Katz*, Chief Justice John Roberts's reasoning displays a more interesting dynamic. Roberts did not rest his opinion so much on a theory of the kinds of spaces in which people should feel secure as on the kinds of subjects we have become in the digital age: in effect, cyborgs. Cell phones, he noted, stand in for many different kinds of recording and storage technologies that register information about private lives: "They could just as easily be called cameras, video players, rolodexes, calendars, tape recorders, li-

braries, diaries, albums, televisions, maps, or newspapers." As such, they are de facto extensions of human selves. Indeed, as the Chief Justice mused, cell phones are "now such a pervasive and insistent part of daily life that the proverbial visitor from Mars might conclude they were an important feature of human anatomy."[40]

It is perhaps not surprising that a judge trained in the common law's style of empirical reasoning imagined a Martian who, like its human counterpart, would focus in the first instance on the visible connections between the cell phone and the human anatomy. Yet the decision turned on a more subtle difference between the cell phone and any other device a person might be carrying. It was the capacity of the phone to provide entry into a person's consciousness – by revealing contacts, photographs, e-mail, telephone data, Internet search records – that was at stake in the ruling. The material object, in other words, makes the normally locked and protected spaces of the human mind visible to prying eyes. To claim a cell phone is "materially indistinguishable" from any other physical object was, Roberts therefore concluded, "like saying a ride on horseback is materially indistinguishable from a flight to the moon."

Looking across the broad terrain of legal encounters with science and technology, it is hard to ignore the extent to which judges in the U.S. legal system have become transmitters of cultural common sense, particularly in their views on the right ways to integrate scientific knowledge and technical expertise into the fabric of the law. Even in those areas where the law explicitly defers to science, as in *Daubert*'s injunction to judges to think like scientists, we find that deference is filtered through preexisting judicial ideas that shape choices at crucial junctures: how the law should accommodate changes in science; who counts as an authoritative expert; and how new objects should be classified for purposes of applying established legal rules.

Despite *Daubert*'s supposedly revolutionary impact on the admissibility of evidence, a close look at *Kumho* shows how quickly judicial common sense reasserted itself, consolidating even greater power over a wider range of knowledge in the hands of the judge. Deeply enmeshed within that expansion of power was an epistemic tilt toward the credibility of the eyewitness above the abstracted, probabilistic knowledge of the witness who appeals to scientific theory. Under the guise of better science in the courtroom and more rigorous assessment of scientific evidence, the law thus reasserted its ancient sources of authority: case-by-case reasoning and the fundamental role of direct eyewitnessing, nominally guided by the *Daubert* criteria as a stronger armature for older forms of judicial empiricism.

Common sense in its nature is unreflective. In Geertz's terms, it steps in as "what everybody knows" and is readily accepted for that very reason. Judicial common sense is no exception: yet there has been little systematic inquiry into how judges think about science and technology, let alone into the consequences of buying into particular theories of the scientific method or technological change. Common sense ensures a kind of stability in the workings of society, and its role in legal reasoning may, in that respect, serve a valuable function as an affirmation of important communal norms and a safeguard against overly rapid and arbitrary turns of the wheel. Yet when federal judges serve society over many decades, one may ask whether such lack of self-awareness in the law is an unmitigated public good. More than having judges think like scientists, both the judiciary and society would benefit from deeper reflection on what it means – in societies transformed by scientific and technological change – to think like judges about science, evidence, and invention.

AUTHOR'S NOTE

This essay draws on research conducted under a grant from the National Science Foundation (Award No. 0850962). I would like to thank Ellen Bales for significant background research on the Supreme Court's evidence decisions and collaboration on a working paper that served as a partial basis for this essay.

ENDNOTES

[1] Tal Golan, *Laws of Men and Laws of Nature: The History of Scientific Expert Testimony in England and America* (Cambridge, Mass.: Harvard University Press, 2004).

[2] Clifford Geertz, "Common Sense as a Cultural System," *The Antioch Review* 33 (1) (1975): 788.

[3] Ibid., 781.

[4] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

[5] *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923).

[6] As elaborated in *Daubert*, the Rules require a trial court to take into account a number of considerations, with special attention to the ones mentioned below, that might affect the reliability of expert scientific evidence.

[7] Sheila Jasanoff, "Law's Knowledge: Science for Justice in Legal Settings," *American Journal of Public Health* 95 (S1) (2005): S49 – S58.

[8] *Kumho Tire Company, Ltd., et al. v. Patrick Carmichael et al.*, 526 U.S. 137 (1999).

[9] See also Sheila Jasanoff, "Science and the Statistical Victim: Modernizing Knowledge in Breast Implant Litigation," *Social Studies of Science* 32 (1) (2002): 37 – 69.

[10] *Patrick Carmichael et al. v. Samyang Tire, Inc., et al.*, 131 F.3d 1433, 1435 (11th Cir. 1997).

[11] *Berry v. City of Detroit*, 25 F.3d 1342 (6th Cir.1994).

[12] Ibid., 1350, quoting *Patrick Carmichael et al. v. Samyang Tire, Inc., et al.*, 131 F.3d, 1435 – 1436.

[13] *Patrick Carmichael et al. v. Samyang Tire, Inc., et al.*, 131 F.3d 1435 (11th Cir. 1997).

[14] Ibid.

[15] Steven Shapin and Barry Barnes, "Head and Hand: Rhetorical Resources in British Pedagogical Writing, 1770 – 1850," *Oxford Review of Education* 2 (3) (1976): 231 – 254.

[16] Oral Arguments, *Kumho Tire Company, Ltd., et al. v. Patrick Carmichael et al.*, 1998 U.S. Trans. Lexis 80, 42 – 43.

[17] Author conversation with Jed S. Rakoff, *Dædalus* Authors' Conference, Science and the Legal System, Cambridge, Massachusetts, July 20 – 21, 2017. Emphasis added.

[18] Oral Arguments, *Kumho Tire Company, Ltd., et al. v. Patrick Carmichael et al.*, 45 [see note 16]. Emphasis added.

[19] *Kumho Tire Company, Ltd., et al. v. Patrick Carmichael et al.*, 526 U.S., 148 [see note 8].

[20] Ibid.

[21] Ibid., 155.

[22] Jasanoff, "Science and the Statistical Victim," 49 [see note 9].

[23] *Massachusetts v. Environmental Protection Agency*, 549 U.S. 497 (2007).

[24] Ibid.

[25] Ibid., 559 – 560.

[26] Textualism of the kind Justice Scalia adopted in his dissent is more of a strategy of persuasion than a definitive pinning down of what a legal text really means. The meaning a judge ac-

cords to legal language is conditioned in any case by prior cultural and personal understandings of which words to question and what those words can be made to mean in context. The important point here, however, is that in resorting to the dictionary as the definitive, disambiguating authority on the meaning of the Clean Air Act, Justice Scalia sought to sidestep the majority's reliance on scientific consensus as a basis for rereading the law in new, more expansive ways. As long as common-sense dictionary definitions supported a reading that *could* justify EPA's inaction, the right recourse, in Scalia's scheme of things, would have been to return to Congress to clarify the law, focusing in this case not on the arcane details of climate science but on the plain meaning of the word "air."

[27] *Diamond v. Chakrabarty*, 447 U.S. 303 (1980).

[28] Ibid., 317.

[29] *Association for Molecular Pathology v. Myriad Genetics, Inc.*, 569 U.S. 576 (2013).

[30] *Association for Molecular Pathology v. U.S. Patent & Trademark Office*, 689 F.3d 1303, 1326 (CAFC 2012).

[31] *Association for Molecular Pathology v. Myriad Genetics, Inc.*, oral argument transcript, 5, https://patentlyo.com/media/docs/2013/04/12-398-amc7.pdf (accessed July 2018).

[32] *Association for Molecular Pathology v. Myriad Genetics, Inc.*, 569 U.S. 576, 591 [see note 29]. See also Chris Hansen's opening argument: "Myriad unlocked the secrets of two human genes. These are genes that correlate with an increased risk of breast or ovarian cancer. But the genes themselves, their – where they start and stop, what they do, what they are made of, and what happens when they go wrong are all decisions that were made by nature, not by Myriad. Now, Myriad deserves credit for having unlocked these secrets. Myriad does not deserve a patent for it." *Association for Molecular Pathology v. Myriad Genetics, Inc.*, oral argument transcript, 3, https://patentlyo.com/media/docs/2013/04/12-398-amc7.pdf (accessed July 2018) [see note 31].

[33] *Association for Molecular Pathology v. Myriad Genetics, Inc.*, 569 U.S. 576, 591 [see note 29].

[34] Ibid., 590.

[35] *Katz v. United States*, 389 U.S. 347 (1967).

[36] Ibid., 360.

[37] Arguably, this is the principle to be extracted from the agreement reached between the National Institutes of Health (NIH) and the family of Henrietta Lacks, whose dying body was the source of the HeLa cell line, although the NIH declared the arrangement to be *sui generis* and of no precedential value.

[38] Case C-131/12, *Google Spain SL v. Agencia Española de Protección de Datos* (May 13, 2014).

[39] *Riley v. California*, 573 U.S. ___ (2014).

[40] Ibid.

The Supreme Court & Science: A Case in Point

Linda Greenhouse

Abstract: When it comes to science and technology, Supreme Court justices resemble lay people in robes, often ill-equipped to grasp fully the implications of the important cases they are asked to decide on scientific subjects. The justices approach science not in the abstract, of course, but from within the doctrinal area in which the particular dispute arises, whether intellectual property, criminal law, or the First Amendment's protection of free speech. The Supreme Court's abortion jurisprudence offers a particularly interesting and consequential example of the Court's encounter with science: a prolonged encounter, since from the beginning, the Court viewed women's claim to reproductive freedom through a medicalized lens. In recent years, states wishing to curb access to abortion have claimed health justifications for placing novel and onerous restrictions on abortion providers. In Whole Woman's Health v. Hellerstedt, *decided in June 2016, the Court invalidated one such effort, a Texas law, on the ground that the claimed health benefits were insufficient to justify the predictably massive shrinkage of the medical infrastructure necessary for women to be able to exercise their constitutional right to terminate a pregnancy. Evidence-based law met evidence-based medicine in a decision that demonstrated a new willingness by the Court to insist on good science in the area of abortion, and perhaps beyond.*

LINDA GREENHOUSE, a Fellow of the American Academy since 1994 and a member of the Academy's Council and its Committee on Studies and Publications, is the Joseph Goldstein Lecturer in Law at Yale Law School. She is also President of the American Philosophical Society. She covered the Supreme Court for *The New York Times* from 1978 to 2008 and continues to write a biweekly column on law as a contributing columnist for *The New York Times* website. She is the author, most recently, of *Just a Journalist: On the Press, Life, and the Spaces Between* (2017).

Science and the Supreme Court of the United States are uneasy partners. Justice Antonin Scalia made that quite clear in a one-paragraph opinion concurring in the Court's unanimous 2013 decision on the patentability of sequences in the human genome: in this case, genetic mutations that increase the risk of breast and ovarian cancer. "I join the judgment of the Court, and all of its opinion" except for those sections describing "fine details of molecular biology," Justice Scalia wrote in *Association for Molecular Pathology v. Myriad Genetics*. He explained: "I am unable to affirm those details on my own knowledge or even my own belief."[1]

This was surely an odd expression of insecurity from the ordinarily self-confident justice. What sort of "belief" in molecular biology was he lacking? (Or,

doi:10.1162/DAED_a_00518

by the same token, on what beliefs regarding other subjects on the Supreme Court's docket was he content to rely without question?)

Justice Scalia was no longer alive when, during its 2016 term, the Court considered the question of how courts should measure intellectual disability, for purposes of deciding whether a capital defendant should be deemed so disabled as to be constitutionally ineligible to be put to death. The Texas Court of Criminal Appeals had rejected the definitional approach to intellectual disability currently used in the medical community. Upholding the death sentence for a man with IQ scores in the 70s and adaptive-functioning test scores more than two standard deviations below the mean, the state court instead employed a guideline from a 1992 opinion of its own. As Justice Ruth Bader Ginsburg described the inadequacy of that measure in her majority opinion, it relied on "lay perceptions of intellectual disability" long superseded by "improved understanding over time."[2] The Supreme Court overturned the death sentence. Justice Ginsburg's opinion canvased the current medical approach, relying in part on a brief filed by the American Psychological Association that described contemporary understanding and practice.

Notably, the decision in *Moore v. Texas* was not unanimous. In his dissenting opinion, which Justices Clarence Thomas and Samuel Alito joined, Chief Justice John Roberts objected that the definition of "cruel and unusual" punishment – a punishment that thus violates the Eighth Amendment – must rest "on a judicial judgment about societal standards of decency, not a medical assessment of clinical practice." The chief justice continued: "The Eighth Amendment, under our precedent, is supposed to impose a moral backstop on punishment, prohibiting sentences that our society deems repugnant. The Court, however, interprets that constitutional guarantee as turning on clinical guidelines that do not purport to reflect standards of decency."[3]

This was a fascinating objection: not that the current medical standards were incorrect or incapable of consistent application, but that outside a particular constitutional context, they were simply irrelevant. The dispute in this Texas death penalty case thus has profound implications across the Supreme Court's docket, whenever the justices are faced with deciding what weight to give a claim based on science compared with the weight of a claim grounded in precedent or in the deference owed to Congress or a state legislature.

In other words, a Supreme Court case is not a laboratory experiment, and science does not reside on the Court's docket in a vacuum.[4] It always exists in context. And the most freighted context of all is abortion.

"In the abortion area," one scholar of abortion law observed not long ago, "law drives science more than science drives law."[5] While that statement may appear paradoxical, it simply reflects the framework judges use to rule on constitutional questions. The relevance of science – or history, or economics, or any field of knowledge extrinsic to the actual legal materials at hand – turns on how closely judges are prepared to scrutinize the legislation they are reviewing. The degree of judicial scrutiny determines how much deference courts give to legislative actions. Thus, a law that touches on purely economic interests, which under the Supreme Court's precedents receive the lowest level of judicial scrutiny, will ordinarily be upheld as long as judges are satisfied that there was some reason, almost any reason, for its enactment.[6] Under "rational basis review," judicial deference to legislative choice is nearly total. At the other end of the spectrum, the government needs a "compelling" justification for infringing a right deemed "fundamental," paradigmat-

ically the right to be free of official discrimination on the basis of race.[7] Such laws or government policies are accorded strict judicial scrutiny, with courts' deference at a minimum.

With that brief digression into constitutional law as background, I turn now to a slightly more extended survey of the Supreme Court's abortion jurisprudence. What emerges – and what is too often overlooked in discussions about the Court and abortion – is the extent to which law and medicine intersect and entwine, from the beginning of the story through the Court's most recent decision.[8]

To begin at the beginning: The Court's 1973 decision in *Roe v. Wade* recognized as "fundamental" a woman's right to terminate a pregnancy before fetal viability. After viability, according to *Roe*, the state acquires a "compelling" interest in unborn life and can prohibit abortion except when necessary to preserve a woman's life or health.[9] The Court explained that

> for the period of pregnancy prior to this "compelling" point, the attending physician, in consultation with his patient, is free to determine, without regulation by the State, that, in his medical judgment, the patient's pregnancy should be terminated. If that decision is reached, the judgment may be effectuated by an abortion free of interference by the State.[10]

Note that the Court, far from hoisting a banner for women's rights, placed the decision of whether to terminate a pregnancy in the hands of the (presumably male) doctor "in consultation with his patient," rather than, as one might suppose, the other way around. *Roe v. Wade* was a highly medicalized decision, relying both on a medical definition of the course of a pregnancy and on doctors to make the appropriate decision.[11] While a feminist expression of the abortion right might declare that it is not up to the state to determine a woman's life course, the Court's primary concern lay elsewhere. The men who voted with the seven-to-two majority to recognize a right to abortion were concerned with protecting their peers in the medical profession against criminal prosecution for applying their best judgment of how to deal with a patient's undesired or compromised pregnancy. It was not the role of the state, the Court declared, to second-guess the exercise of professional judgment.[12]

For some years after *Roe*, with the majority supporting the right to abortion still largely intact, the Court adhered to the view that a doctor's judgment was not to be questioned. For example, in the 1979 case *Colautti v. Franklin*, the Court struck down a Pennsylvania law that required doctors to perform later-term abortions by the method most likely to preserve the life of a potentially viable fetus. The six justices in the majority recoiled from the notion that Pennsylvania was telling doctors what to do, on pain of criminal liability, in such a delicate and often ambiguous situation.[13] "The choice of an appropriate abortion technique," Justice Harry Blackmun wrote for the Court, involved "a complex medical judgment about which experts can – and do – disagree." Clearly, this was a matter for doctors, not legislators.

But then Ronald Reagan was elected president, on a platform that called for overturning *Roe v. Wade*, and things began to change. Justice Potter Stewart, a strong member of the *Roe* majority, retired in 1981 and was succeeded by Sandra Day O'Connor. Her views on *Roe v. Wade* remained a mystery for her first two years on the bench. But near the end of her second term, the Court decided a case from Akron, Ohio. The city had enacted an ordinance entitled "Regulation of Abortion" that, among other features, imposed a twenty-four-hour waiting period and required doctors to read an "informed consent" script that the city fathers hoped might persuade women to change

their minds. The Court invalidated the ordinance, with Justice Powell explaining for the majority that fidelity to *Roe v. Wade* left no choice but to declare the ordinance unconstitutional.[14]

Justice O'Connor dissented in a strongly worded attack on *Roe* itself, centered on her understanding of neonatology. Premature infants were being saved at ever earlier gestational ages, she wrote, observing that a baby born at twenty-two weeks "is now thriving in a Los Angeles hospital."[15] She continued:

> It is certainly reasonable to believe that fetal viability in the first trimester of pregnancy may be possible in the not too distant future. . . . The *Roe* framework, then, is clearly on a collision course with itself. As the medical risks of various abortion procedures decrease, the point at which the State may regulate for reasons of maternal health is moved further forward to actual childbirth. As medical science becomes better able to provide for the separate existence of the fetus, the point of viability is moved further back toward conception.[16]

This was a powerful critique, invoking a medical framework to attack the core of the medicalized *Roe* itself. It was, however, not accurate. While during the decade since *Roe* the survival rate for extremely premature early third-trimester infants like the one Justice O'Connor described had improved from 2 percent to about 10 percent, that did not mean that viability was moving back through the second and first trimesters toward conception. When a case presenting a frontal attack on *Roe* appeared on the Supreme Court's docket six years later, the medical community mobilized to make sure the science of gestation and neonatology would be clear to the justices.

That opportunity came in a case from Missouri, *Webster v. Reproductive Health Services, Inc*.[17] Justice O'Connor's overt hostility was not the only development that had placed the future of *Roe v. Wade* in grave doubt. So had Justice Powell's recent retirement and his replacement by Anthony M. Kennedy. Further, the administration of President George H. W. Bush raised the stakes by entering the case as a "friend of the Court" to argue vigorously for *Roe*'s overruling. *Amicus curiae* briefs flooded into the Court – seventy-eight of them, a record at the time.[18] For our purposes, the most directly relevant was a brief filed by a coalition of professional medical organizations that included the American Medical Association, the American College of Obstetricians and Gynecologists, and the American Academy of Pediatrics. This brief described the existence of an "anatomic threshold" at twenty-three to twenty-four weeks of gestation; earlier than that, it explained, "the fetal lung does not mature sufficiently to permit normal or even mechanically-assisted respiration."[19] The brief added that medical intervention before that point was fruitless and that "improvements are not expected in the foreseeable future."[20]

When *Webster* was decided on July 3, 1989, Justice O'Connor refused to join the four justices who would either have overruled *Roe* explicitly, as Justice Scalia advocated, or would have relegated the right to abortion to mere rational-basis review, the position taken by Chief Justice William Rehnquist and Justices Byron White and Kennedy. While voting to uphold the particular regulations at issue, Justice O'Connor said there was no need to revisit *Roe* itself: "When the constitutional invalidity of a State's abortion statute actually turns on the constitutional validity of *Roe*, there will be time enough to reexamine *Roe*, and to do so carefully."[21] Her separate opinion did not refer to *Roe*'s purported "collision course with itself." Had she read the medical brief and become persuaded that her instinctive conclusion about the future course of viability was scientifically un-

sound? The only evidence we have to go on is the fact that she never mentioned the collision course again.

The Supreme Court's next opportunity to overturn *Roe v. Wade* came only three years later, in *Planned Parenthood of Southeast Pennsylvania v. Casey*.[22] Much had changed, and *Roe*'s prospects appeared even more dire: two more members of the original *Roe* majority, Justices Brennan and Marshall, had retired, and President George H. W. Bush had replaced them with Justices David Souter and Thomas. But the Court surprised nearly everyone by reaffirming the right to abortion by a vote of five to four, with Justices O'Connor, Kennedy, and Souter producing an unusual joint opinion that announced a new approach to evaluating abortion regulations: the undue-burden standard.

First proposed by Justice O'Connor in her *Akron* dissent, the undue-burden standard remains the law of the land today. The *Casey* decision defined an undue burden as "a state regulation [that] has the purpose or effect of placing a substantial obstacle in the path of a woman seeking an abortion of a nonviable fetus."[23] While those words were clear enough, their application was anything but certain. It was evident that the fundamental-rights language of *Roe*, with its implication of strict judicial scrutiny of any obstacle to access to abortion, had been superseded. But what did this mean in practice? What type of obstacle was "substantial"? What level of judicial scrutiny was now required?

Rather than answer those questions explicitly, the Court proceeded by example. In *Casey* itself, it upheld most of the challenged regulations contained in Pennsylvania's Abortion Control Act of 1982, including the same waiting-period and mandatory-counseling requirements that had been declared unconstitutional nine years earlier in the *Akron* case.[24] At the same time, the Court struck down as an undue burden a requirement that a married woman inform her husband of her plan to terminate a pregnancy. The major regulations addressed by the Court in *Casey* thus concerned the state's ability to dissuade a woman from terminating her pregnancy. None directly concerned women's health, so one sentence nearly fifty pages into the principal opinion seemed almost beside the point at the time, attracting little notice: "Unnecessary health regulations that have the purpose or effect of presenting a substantial obstacle to a woman seeking an abortion impose an undue burden on the right."[25]

Fifteen years after *Casey*, in 2007, with Justice O'Connor having been succeeded by Justice Alito, the Court upheld the federal Partial Birth Abortion Ban Act, a law that made it a crime for a doctor to use an abortion method known medically as "intact dilation and extraction" and made notorious by abortion opponents under the label they gave it, "partial-birth abortion."[26] The undue-burden question for the Court in this case, *Gonzales v. Carhart*, was whether the procedure was ever medically necessary, given the availability of more common methods of second-trimester abortion. (And if the procedure was regarded as medically necessary, the law would have to provide for an exception from the criminal ban when a woman's health or life was at stake.) Finding a division of medical opinion on the question – as established by extensive district court litigation in the case – the Court deferred to the congressional judgment that no exception to the ban was required; the absence of a health exception therefore did not amount to an undue burden.[27] At the same time, Justice Kennedy made it clear in the majority opinion that the Court's deference to Congress was neither automatic nor complete. His language, although little noticed at the time, would prove sig-

nificant: "The Court retains an independent constitutional duty to review factual findings where constitutional rights are at stake.... Uncritical deference to Congress's factual findings in these cases is inappropriate."[28]

While this was the Supreme Court's last word on the meaning of undue burden for a decade, the abortion landscape outside the Court was hardly quiescent. Abortion opponents, frustrated by the failure of frontal attacks on the right to abortion itself, shifted their focus to the clinic infrastructure necessary to keep abortion relatively accessible and affordable. Spurred by the effective advocacy of Americans United for Life (AUL), a well-established generator of abortion-restricting legislative proposals, Republican-dominated states began to enact laws with the ostensible goal of protecting women from "an increasingly under-regulated and rapacious abortion industry," in the words of Americans United for Life's 471-page handbook *Defending Life*.[29] Many of the legislatures enacting these laws have followed templates provided by AUL's "Women's Protection Project": "strategic, life-affirming legislation" described as protecting women from "abortion industry abuse."[30] Among these were laws requiring doctors who perform abortions to have admitting privileges at nearby hospitals, and requiring the clinics themselves to meet the physical and operational standards required of ambulatory surgical centers. The abortion-rights community labeled these statutes TRAP laws, for "targeted regulation of abortion providers," underscoring the fact that abortion was being singled out and that no similar requirements were imposed on providers of medical services considerably riskier than abortion, including liposuction, colonoscopy, and arthroscopic surgery. In earlier work, Reva B. Siegel and I referred to these abortion-targeting laws as manifestations of "abortion exceptionalism."[31] Others have referred to such laws as "supply-side policies."[32]

The Texas law known as H.B. 2 – enacted in 2013 – imposed both the admitting privileges and the ambulatory surgical center requirements. At the time, there were forty-two abortion clinics in Texas. The state had long required abortion practices to maintain transfer agreements with outside doctors who would be available to care for any patients needing hospitalization. But it had not required clinic doctors themselves to have admitting privileges, and in eighteen of the forty-two clinics, there were no doctors who had them. And only six clinics, all located in four major cities (Austin, Fort Worth, Houston, and San Antonio), met the surgical-center requirement.

The abortion clinics went immediately to federal court to challenge the constitutionality of the new requirements. How would the courts respond? The sponsors of H.B. 2, following the Americans United for Life playbook, presented the law as necessary to protect the health of Texas women. The legislators were clearly aware that the law would close clinics, and even which specific clinics would be affected. The day after the bill cleared the state Senate (where it was known as S.B. 5), David Dewhurst, the lieutenant governor at the time, tweeted a picture of a map showing the clinics that would close and exulted: "We fought to pass S.B. 5 last night, & this is why!"[33]

Challenges to the surgical-center provision and the admitting-privileges requirement were litigated separately. Each reached Federal District Judge Lee Yeakel of the United States District Court in Austin. In October 2013, Judge Yeakel enjoined the admitting privileges requirement, finding that it bore "no rational relationship to improved patient care" or to "the State's legitimate interest in protecting the unborn."[34] He elaborated: The hospital committees that confer admitting privileges typically require a number of patient ad-

missions each year. But so few abortion patients ever needed hospitalization that doctors whose practice consisted largely of abortions were unable to meet the quota. Judge Yeakel emphasized that from the perspective of patient care, there was no cause for concern; he quoted trial testimony from an emergency room doctor who said that there would be no difference in treatment for an abortion patient regardless of whether her doctor had admitting privileges or lacked them. The state not only "fails to show a valid purpose for the requirement," Judge Yeakel continued, but "the evidence is that clinics will close" as a result. The admitting privileges requirement, he concluded, thus imposed an undue burden on the right to abortion.[35]

Clinics did close, nearly half of all the abortion clinics in Texas, after the United States Court of Appeals for the Fifth Circuit overturned Judge Yeakel's injunction and then refused to issue a stay of its ruling to enable the clinics to appeal to the Supreme Court.[36] For our purposes, what was notable about the appeals court's ruling was its approach to the facts of the case. Did the legislature's asserted health justification for the admitting privileges requirement hold up to inspection? The Fifth Circuit offered no conclusion because, the court said firmly, the answer to that question did not matter. Abortion regulations were subject only to "*rational* basis review, not *empirical* basis review," the court said.[37] This highly deferential test, the opinion went on, "affirms a vital principle of democratic self-government" and "seeks only to determine whether any conceivable rationale exists for an enactment."[38]

This was just the beginning. After Judge Yeakel, in a subsequent opinion, struck down the ambulatory surgical center requirement,[39] the Fifth Circuit not only overturned his decision but rebuked him for even "evaluat[ing] whether the ambulatory surgical center provision would actually improve women's health and safety." The court emphasized: "In our circuit, we do not balance the wisdom or effectiveness of a law against the burdens the law imposes."[40] This was a swipe at a recent decision by another federal appeals court, the Seventh Circuit, blocking enforcement of an admitting privileges law in Wisconsin. Writing for that court, Judge Richard A. Posner had noted with evident exasperation that despite the asserted health-protecting purpose for requiring admitting privileges, "no documentation of medical need for such a requirement was presented to the Wisconsin legislature."[41] Judge Posner observed that while the requirement would shut half the state's abortion clinics, the medical evidence for it was "feeble" at best. He interpreted the undue-burden standard to require a kind of weighted balancing test: "The feebler the medical grounds, the likelier the burden, even if slight, to be 'undue' in the sense of disproportionate or gratuitous."[42]

The issue was joined. Did medical or scientific evidence matter to the law of abortion, or did it not? The Fifth Circuit's invocation of a rational basis test, one so deferential that a trial judge was obliged to ignore pertinent evidence, appeared to be flatly incorrect. After all, in adopting the undue-burden standard, the Court in *Casey* rejected the argument that a rational-basis test was constitutionally sufficient; those justices who argued for rational basis did so in dissent.[43] But *Casey* was a generation ago, and some viewed the Roberts Court's intervening *Gonzales v. Carhart* decision as having lowered the standard to something close to rational basis (a conclusion that required overlooking Justice Kennedy's admonition in that case that "the Court retains an independent constitutional duty to review factual findings where constitutional rights are at stake").[44] On November 13, 2015, the Supreme Court agreed to hear the clinics' appeal of the Fifth Cir-

cuit's decision. What the decision would be was anyone's guess.

Issued on June 27, 2016, the Court's decision in *Whole Woman's Health v. Hellerstedt* invalidated both requirements of the Texas statute.[45] And it did much more. It reanimated the undue-burden standard, making clear that the appeals court had been mistaken in its unquestioning deference to the legislature's health claims. Judge Yeakel had been correct to test those claims against the medical evidence available, Justice Stephen Breyer wrote for the five-to-three majority. "For a district court to give significant weight to evidence in the judicial record in these circumstances is consistent with this Court's case law," Justice Breyer said. He explained that, contrary to the Fifth Circuit's complaint, the District Court

> did not simply substitute its own judgment for that of the legislature. It considered the evidence in the record – including expert evidence presented in stipulations, depositions, and testimony. It then weighed the asserted benefits against the burdens. We hold that, in so doing, the District Court applied the correct legal standard.[46]

With a minimum of rhetoric – there are no ringing phrases in Justice Breyer's twenty-page opinion – but a plethora of facts, the Court demolished the state's justification for its clinic-closing law. On the benefit side of the benefit-versus-burden equation, Justice Breyer recounted the evidence Judge Yeakel had compiled about the safety record for abortion in Texas, concluding that "there was no significant health-related problem that the new law helped to cure."[47] Without labeling the law as abortion exceptionalism, he noted that although abortion is fourteen times safer than childbirth, Texas "allows a midwife to oversee childbirth in the patient's own home," and that while liposuction has a twenty-eight times higher mortality rate than abortion, there are no similar surgical-center requirements for performing that procedure on an outpatient basis.[48]

Reviewing the evidence underlying the admitting privileges requirement, Justice Breyer said that "without dispute," the basis on which admitting privileges are granted in the context of abortion has "nothing to do with ability to perform medical procedures" and "does not serve any relevant credentialing function." There was a "virtual absence of any health benefit," he said in recalling one of the most dramatic moments of the March 2, 2016, oral argument: "When directly asked at oral argument whether Texas knew of a single instance in which the new requirement would have helped even one woman obtain better treatment, Texas admitted that there was no evidence in the record of such a case."[49]

Having dispensed with the health justification, the Court then turned to the burden the new requirements had already imposed on the clinics and would predictably impose on Texas women's access to abortion. Justice Breyer noted that the closing of half of the state's abortion clinics, with the imminent prospect of more closings once the surgical-center requirement went into effect, "meant fewer doctors, longer waiting times, and increased crowding," along with more than quadrupling, to four hundred thousand, the number of women of reproductive age living more than 150 miles from an abortion provider.[50] He said that "in the face of no threat to women's health, Texas seeks to force women to travel long distances to get abortion in crammed-to-capacity superfacilities. Patients seeking these services are less likely to get the kind of individualized attention, serious conversation, and emotional support that doctors at less taxed facilities may have offered." It was a "commonsense inference," Justice Breyer concluded, "that

these effects would be harmful to, not supportive of, women's health."[51]

The decision was cheered in the medical community. An article in *Obstetrics & Gynecology* declared that the decision's analytical framework

> recalibrates the debate over abortion laws from one that has too often been mired in rancor and rhetoric to one that is rooted in data and science. . . . With *Hellerstedt*, the Supreme Court has not only "talked the talk" about the importance of evidence, but has "walked the walk" by allowing that evidence to drive its analysis.[52]

The Court's appreciation of the impact of abortion restrictions on the lives of actual women is a distinctive feature of the decision.[53] To that extent, *Whole Woman's Health* is abortion-specific. The decision is likely to prove useful in attacking other scientifically unsupported abortion restrictions. One example is the prohibition adopted in some states against the use of telemedicine for dispensing the pills prescribed for terminating first-trimester pregnancies. Another are the bans that states are now imposing on abortion beginning at twenty weeks of pregnancy, based on the unsupported claim that a fetus, while not viable at that gestational age, feels pain.[54] The decision may also be useful in challenging mandatory counseling laws that require doctors to give women false information about the consequences of abortion, such as warning that abortion increases the risk of breast cancer and suicide. Both those claims have been extensively studied and refuted.[55]

But whether *Whole Woman's Health* may help in challenging another category of abortion restrictions – those adopted not in the name of protecting women, but rather to express the state's interest in protecting unborn life – remains an open question. One example is a Texas law enacted in 2017 to require fetal remains obtained through abortion (although not through miscarriage) to be cremated or buried. That this law will serve to increase the cost of abortion is clear, although the means for attacking the law are less so.[56] By personifying the fetus, the law is also likely, not coincidentally, to increase the stigma attached to abortion, a burden already felt by women who choose to terminate a pregnancy. Research has shown that most women try to keep their abortions secret out of concern for how even close friends and family would respond.[57] Texas describes its motive as a desire to express the state's view of the dignity of unborn life, a state interest that the Supreme Court's abortion jurisprudence protects. With the case now being litigated, it remains to be seen how the revivified undue-burden analysis of *Whole Woman's Health* will apply in this context. There is no reason it should not. The undue-burden standard itself derives from *Casey*, which applied it to regulations explicitly aimed at protecting unborn life.

It is nonetheless evident that legislatures and courts with antiabortion majorities are not accepting the lessons of *Whole Woman's Health* without protest.[58] Arkansas is defending a 2015 law that requires doctors who provide medication abortions – the abortion-inducing drugs administered to terminate early pregnancies – to have a signed contract with a doctor who has admitting privileges at a local hospital in the case of an emergency. The state's claimed rationale is to protect women's health. Medication abortion is extremely safe; fewer than one-third of 1 percent of such abortions result in any adverse event. (Nor is telemedicine, which eighteen states prohibit for medication abortions, any less safe, according to a recent article in *Obstetrics & Gynecology*.)[59] The local Planned Parenthood affiliate testified that it could not find a physician willing to sign a contract, and would therefore have to stop providing medication abortions at its two clinics; it provided no surgi-

cal abortions at those facilities. That would leave only one provider in the state, in Little Rock. A federal district judge enjoined the law in March 2016, finding that the burden on women seeking abortions – which for women living in Fayetteville would include two 380-mile round-trips to Little Rock – outweighed any asserted benefit.[60] In July 2017, the U.S. Court of Appeals for the Eighth Circuit lifted the injunction. The court cited *Whole Woman's Health* without actually applying it, instead finding the district court's analysis of the law's burden too "amorphous" without making any effort to analyze the law's asserted benefit. The district court had "failed to make factual findings estimating the number of women burdened by the statute," the appeals court complained.[61] It is difficult to read the Eighth Circuit's opinion as anything other than a deliberate evasion of the Supreme Court's mandate in *Whole Woman's Health*. Clearly, in the hands of abortion-hostile courts, *Whole Woman's Health* is not the complete answer to legislatures that invoke bad science, or no science at all, in their crusade to cut off women's access to abortion. Planned Parenthood sought Supreme Court review, but on May 29, 2018, the court denied the petition without comment or noted dissent.[62] Under the terms of the Eighth Circuit's order, the case returned to the district court for more factual development. On July 2, 2018, following a new hearing and additional briefing, Federal District Judge Kristine G. Baker issued a new injunction. She found that the law posed "a threat of irreparable harm" to the plaintiffs that "outweighs the immediate interests and potential injury to the state."[63]

Outside the highly politicized context of abortion, it would be reassuring to suppose that *Whole Woman's Health* might strengthen the Supreme Court's resolve to use the legal tools available to separate scientific knowledge from agenda-driven claims that masquerade as science. In *Whole Woman's Health*, evidence-based law met evidence-based medicine in a manner that should serve as a template for judicial encounters with the science and technology that will increasingly shape the world that judges, along with the rest of us, inhabit. Whether it has a chance of filling that role depends on politics and on future appointments to the Court – contingencies that even the best science cannot control.

ENDNOTES

1 *Association for Molecular Pathology v. Myriad Genetics, Inc.*, 133 S. Ct. 2107, 2120 (2013), Antonin Scalia concurring in part and concurring in the judgment. The decision, with an opinion by Justice Clarence Thomas, rejected patentability for naturally occurring DNA but not for the synthetic variant the company had created, known as complementary DNA or cDNA.

2 *Moore v. Texas*, 137 S. Ct. 1039, 1051, 1053 (2017).

3 Ibid., 1058, John Roberts dissenting.

4 Of many recent examples of the Court's struggles with science and technology, one of the more interesting was *City of Ontario v. Quon*, 560 U.S. 746 (2010). In this case, a public employee challenged the employer's right to search the text messages on his office-issued pager. The case attracted considerable notice while it was pending in the expectation that the Court might issue a broad rule on the privacy of people's electronic devices. But the Court instead rejected the employee's claim on narrow grounds, with Justice Kennedy explaining that "[r]apid changes in the dynamics of communication and information transmission are evident not just in the technology itself but in what society accepts as proper behavior. . . . At present, it is uncertain how workplace norms, and the law's treatment of them, will evolve. . . . A broad

holding concerning employee's privacy expectations vis-à-vis employer-provided technological equipment might have implications for future cases that cannot be predicted. It is preferable to dispose of this case on narrower grounds." Ibid., 759 – 760. In 2018, the Court did grapple with the privacy implications of electronic devices, holding that the police generally need a warrant to obtain the moment-by-moment location information that wireless carriers automatically acquire from their customers' smartphones. See *Carpenter v. United States*, 16-402, June 22, 2018.

[5] John A. Robertson, "Science Disputes in Abortion Law," *Texas Law Review* 93 (2015): 1849, 1850.

[6] See, for example, *Williamson v. Lee Optical Co.*, 348 U.S. 483 (1955).

[7] *Regents of the University of California v. Bakke*, 438 U.S. 265 (1978).

[8] For an extended discussion of the role of medicine in the Court's abortion jurisprudence, see Nan D. Hunter, "Justice Blackmun, Abortion, and the Myth of Medical Independence," *Brooklyn Law Review* 72 (2006): 147, 149.

[9] *Roe v. Wade*, 410 U.S. 113, 163 – 164 (1973).

[10] Ibid. The Court held that during the second trimester, when abortion carried greater risk, the state could regulate the procedure for the purpose of protecting the pregnant woman's health.

[11] *Roe*'s author, Justice Harry A. Blackmun, had been general counsel of the Mayo Clinic before becoming a federal judge, and had a lifelong interest in medicine and appreciation for its practice. While working on his opinion, he visited Mayo, in Rochester, Minnesota, where the library staff had compiled for him a file of abortion-related articles. See Linda Greenhouse, *Becoming Justice Blackmun: Harry Blackmun's Supreme Court Journey* (London: Macmillan, 2005), 90 – 91.

[12] Commenting on what she called "the myth of medical independence," Nan D. Hunter observed that "the Justices who decided *Roe* shared a liberal belief in the value of medical authority because they assumed it to be a sphere which could operate independently of the state." Hunter, "Justice Blackmun, Abortion, and the Myth of Medical Independence," 147, 149 [see note 8]. A fact little noticed today is that the main impetus for abortion reform in the late 1950s and early 1960s came not from feminists, but from leaders in the public health community, concerned about the health consequences of illegal abortion, particularly for women without the economic means or sophistication to find safer options. See, for example, Linda Greenhouse and Reva B. Siegel, *Before Roe v. Wade: Voices That Shaped the Abortion Debate Before the Supreme Court's Ruling*, 2nd ed. (New Haven, Conn.: Yale Law Library, 2012), 22 – 29, http://documents.law.yale.edu/before-roe.

[13] *Colautti v. Franklin*, 439 U.S. 379, 401 (1979).

[14] *Akron v. Akron Center for Reproductive Health, Inc.*, 462 U.S. 416 (1983); overruled by *Planned Parenthood of Southeastern Pennsylvania v. Casey*, 505 U.S. 833 (1992).

[15] *Akron v. Akron Center for Reproductive Health, Inc.*, 458, Sandra Day O'Connor dissenting.

[16] Ibid., 457 – 458.

[17] *Webster v. Reproductive Health Services, Inc.*, 492 U.S. 490 (1989).

[18] Kathryn Kolbert, "The *Webster Amicus Curiae* Briefs: Perspectives on the Abortion Controversy and the Role of the Supreme Court," *American Journal of Law and Medicine* 15 (2 and 3) (1989): 153, 154. Major Supreme Court cases now attract more than one hundred *amicus* briefs. The current record was set in the same-sex marriage case, in which 139 were filed; see *Obergefell v. Hodges*, 135 S. Ct. 2071 (2015).

[19] Brief of the American Medical Association et al. as Amici Curiae in Support of Appellees at 7, *Webster v. Reprodctive Health Services, Inc.*, 492 U.S. 490 (1989).

[20] Ibid, 8. Were that brief to be written today, the doctors might have felt compelled to acknowledge recent progress toward creating an artificial womb, remote as the prospect appears that it might actually be of use to humans. See, for example, Rob Stein, "Scientists Create Artificial Womb That Could Help Prematurely Born Babies," NPR All Things Considered, April

25, 2017, http://www.npr.org/sections/health-shots/2017/04/25/525044286/scientists-create-artificial-womb-that-could-help-prematurely-born-babies?sc=17&f=1001.

[21] *Webster v. Reproductive Health Services, Inc.*, 490, 406, Sandra Day O'Connor concurring in part and concurring in the result.

[22] *Planned Parenthood of Southeastern Pennsylvania. v. Casey* 505 U.S. 833 (1992).

[23] Ibid., 877.

[24] The decision overturned the relevant parts of the *Akron* decision and of a subsequent decision, *Thornburgh v. American College of Obstetricians and Gynecologists*, 476 U.S. 747 (1986).

[25] Ibid., 878.

[26] *Gonzales v. Carhart*, 550 U.S. 124 (2007).

[27] Ibid., 143 – 144.

[28] Ibid., 165 – 166.

[29] Americans United for Life, *Defending Life*, 2018 ed. (Washington, D.C.: Americans United for Life, 2018), 18, http://www.aul.org/defending-life-2018. President Trump named Charmaine Yoest, the organization's longtime president and CEO, as an assistant secretary of the Department of Health and Human Services, a position she left after a year.

[30] Ibid., 20.

[31] Linda Greenhouse and Reva B. Siegel, "Casey and the Clinic Closings: When 'Protecting Health' Obstructs Choice," *Yale Law Journal* 125 (5) (2016): 1428, 1446 – 1449.

[32] Scott Cunningham, Jason M. Lindo, Caitlin Myers, and Andrea Schlosser, "How Far is Too Far? New Evidence on Abortion Clinic Closures, Access, and Abortion," National Bureau of Economic Research Working Paper 23366 (Cambridge, Mass.: National Bureau of Economic Research, 2017), 2, http://www.nber.org/papers/w23366.

[33] David Dewhurst (@DavidHDewhurst), Twitter post, June 10, 2013, 7:41 a.m., http://twitter.com/DavidHDewhurst/status/347363442497302528/photo/1, http://perma.cc/3QF2-U6QQ.

[34] *Planned Parenthood of Greater Texas Surgical Health Services. v. Abbott*, 951 F. Supp. 2d 891, 900 (W.D. Tex. 2013).

[35] Ibid.

[36] *Planned Parenthood of Greater Texas Surgical Health Services v. Abbott* [*Abbott I*], 734 F. 3d 406 (5th Cir. 2013). The Supreme Court refused to vacate the appeals court's stay of the district court injunction. In a subsequent opinion, 748 F. 3d 583 (5th Cir. 2014) [*Abbott II*], the Fifth Circuit addressed Judge Yeakel's opinion on the merits and overturned it.

[37] *Abbott II*, 748 F. 3d at 596 [see note 36]; emphasis in original.

[38] Ibid., 594.

[39] *Whole Woman's Health v. Lakey*, 46 F. Supp. 3d 673 (W.D. Tex. 2014).

[40] *Whole Woman's Health v. Lakey*, 769 F. 3d 285, 297 (5th Cir. 2014).

[41] *Planned Parenthood of Wisconsin, Inc. v. Van Hollen*, 738 F. 3d 786, 789 (7th Cir. 2013).

[42] Ibid., 798.

[43] See Greenhouse and Siegel, "Casey and the Clinic Closings," 1435 [see note 31], citing *Casey* at 505 U.S. 966, William Rehnquist concurring in the judgment in part and dissenting in part.

[44] *Gonzales v. Carhart*, 550 U.S. 124, 165 – 166 (2007).

[45] *Whole Woman's Health v. Hellerstedt*, 136 S. Ct. 2292 (2016). Justice Breyer's opinion was joined by Justices Kennedy, Ginsburg, Kagan, and Sotomayor. Justice Alito wrote a dissenting opinion that was joined by Chief Justice Roberts and by Justice Thomas, who also wrote a separate dissenting opinion. Justice Scalia died shortly before the case was argued.

[46] Ibid., 2310.

[47] Ibid., 2311.

[48] Ibid., 2314.

[49] Ibid., 2311 – 2313.

[50] Ibid., 2313.

[51] Ibid., 2318.

[52] Joanne D. Rosen, "Finding Strength in Numbers: The Critical Role of Data in *Whole Woman's Health v. Hellerstedt*," *Obstetrics & Gynecology* 129 (1) (2017): 195, 196.

[53] See Linda Greenhouse and Reva B. Siegel, "The Difference a Whole Woman Makes: Protection for the Abortion Right After *Whole Woman's Health*," *Yale Law Journal* 126 (2016): 149, 161.

[54] Ibid.

[55] See, for example, the published studies discussed in Rachel Benson Gold and Elizabeth Nash, "Flouting the Facts: State Abortion Restrictions Flying in the Face of Science," *Guttmacher Policy Review* 20 (2017): 53, 56. The suicide warning, mandated by a South Dakota law, was upheld by a federal appeals court; see *Planned Parenthood v. Rounds*, 686 F. 3d 998 (8th Cir. 2012). A 2018 study published online in the *American Journal of Psychiatry* followed women for five years after they either received or were denied an abortion to see whether women in either group were at greater risk of suicide. The author found no greater risk for either, and observed that "For women having an abortion, we found that the proportion with any symptoms did not increase but rather decreased over the 5-year period." The article concluded: "Thus, policies requiring that women be warned that they are at increased risk of becoming suicidal if they choose abortion are not evidence based." M. Antonia Biggs, "Five-Year Suicidal Ideation Trajectories Among Women Receiving or Being Denied an Abortion," *American Journal of Psychiatry* Online, May 24, 2018, https://doi.org/10.1176/appi.ajp.2018.18010091.

[56] The law has been enjoined on a pre-enforcement challenge in Federal District Court. *Whole Woman's Health v. Hellerstedt*, No. A-16-CA-1300-SS (W.D. Tex. 2017). The state is appealing. Marissa Evans, "Federal Court Blocks Texas Fetal Remains Burial Rule," *Texas Tribune*, January 27, 2017, https://www.texastribune.org/2017/01/27/fetal-remains-ruling/.

[57] See, for example, The University of Chicago, "Accessing Abortion in Illinois: A Guide for Healthcare and Social Service Providers – Understanding Abortion Stigma and Shame," https://abguide.uchicago.edu/page/understanding-abortion-stigma-and-shame (accessed December 7, 2017).

[58] For a compilation of lower court cases as of November 2017, see Leah M. Litman, "Unduly Burdening Women's Health: How Lower Courts Are Undermining *Whole Woman's Health v. Hellerstedt*," *Michigan Law Review Online* 116 (2017): 50.

[59] Daniel Grossman and Kate Grindlay, "Safety of Medical Abortion Provided through Telemedicine Compared with In Person," *Obstetrics & Gynecology* 130 (4) (2017): 778 – 782.

[60] *Planned Parenthood of Arkansas & Eastern Oklahoma v. Jegley* (E.D. Ark.), 2016 WL 6211310.

[61] *Planned Parenthood of Arkansas & Eastern Oklahoma v. Jegley*, 864 F. 3d 953 (8th Cir., 2017).

[62] *Planned Parenthood of Arkansas & Eastern Oklahoma v. Jegley*, 17-935, cert. den. May 29, 2018.

[63] *Planned Parenthood of Arkansas & Eastern Oklahoma v. Jegley*, case 4:15-cv-00784-KGB (July 2, 2018) at 147.

When Law Calls, Does Science Answer? A Survey of Distinguished Scientists & Engineers

Shari Seidman Diamond & Richard O. Lempert

Abstract: Sound legal decision-making frequently requires the assistance of scientists and engineers. The survey we conducted with the cooperation of the American Academy examines the views of the legal system held by some of the nation's most distinguished scientists and engineers, what motivates them to participate or to refuse to assist in lawsuits when asked, and their assessment of their experiences when they do participate. The survey reveals that a majority of the responding scientists and engineers will agree to participate when asked, and when they turn down requests, the most common reasons are lack of time and absence of relevant expertise. Dissatisfaction with legal procedures is also a deterrent, but our respondents indicated that some procedural changes would make their participation more likely. In addition, participation appears to be associated with a greater belief in the ability of the legal system to deal well with scientific matters.

SHARI SEIDMAN DIAMOND, a Fellow of the American Academy since 2012, is the Howard J. Triens Professor of Law and Professor of Psychology at the Northwestern University Pritzker School of Law, and Research Professor at the American Bar Foundation.

RICHARD O. LEMPERT, a Fellow of the American Academy since 1993, is the Eric Stein Distinguished University Professor of Law and Sociology, emeritus, at the University of Michigan.

(Complete author biographies appear at the end of the essay.)

Sound legal decision-making increasingly depends on sound science. Yet we know remarkably little about how scientists and engineers view the legal system or what leads them to decide whether and how to interact with it. Some commentary indicates that scientists regard the legal system with suspicion and discomfort, but the supporting evidence is largely anecdotal. As a result, it is hard to gauge how deep or widespread these reactions are, and – to the extent they exist – whether they are fueled by accurate information or false impressions. Getting a better handle on relationships between scientists and the law matters because the importance of science for law cannot be disputed.

Ideally, courts and litigants would be able to call on knowledgeable, unbiased scientists and engineers whenever the fair resolution of legal disputes depended on scientific or technical information. The

doi:10.1162/DAED_a_00519

importance of the science-law relationship led us, with the cooperation of the American Academy of Arts and Sciences, to conduct a survey of the Academy's science and engineering members with the goal of providing empirical grounding for discussions about how scientists relate to law. Our survey probes scientists' views of and expert involvement with the legal system, especially as it pertains to involvement in litigation, barriers to involvement, and legal or policy changes that might make scientists more willing to aid courts and lawyers when called upon.

The legal system has long recognized the value of scientific knowledge, and lawyers and judges have sought to make use of it, even while struggling to make sense of what science has to offer. The frustration is poignantly reflected in the words of Judge Baron Hatsell in 1699, when in a homicide trial he spoke to the jury about conflicting expert testimony on the cause of death of a young woman whose body was recovered from a lake:

> The Doctors and Surgeons have talkt a great deal to this purpose, and of the waters going into the Lungs or the Thorax, but unless you have more skill in Anatomy than I, you won't be much edified by it. I acknowledge I never studied Anatomy but I perceive that the Doctors do differ in their Notions about these things.[1]

Scientists, for different reasons, have their own difficulties with how the law goes about its business. As one of our respondents put it:

> Science is about truth. The legal system is about spinning, distorting or suppressing the truth in order to win. The ethos of the two fields is fundamentally different. Even judges are biased and not objective. For these reasons, participation in the legal system is very frustrating for a scientist.

The challenge for the modern American legal system is obvious and increasing, as the frequency and complexity of encounters between science and law have multiplied with the dramatic expansion of legally relevant scientific knowledge. Courts and scientific societies have struggled with the tensions that exist.

Justice Stephen Breyer wrote in 1998 that the law "increasingly requires access to sound science."[2] Citing examples of cases on the U.S. Supreme Court's docket, he identified a range of difficult legal problems that implicated scientific, medical, and engineering questions. In lower courts too, both civil and criminal, scientific claims, along with arguments about the quality of expert testimony, are expanding features of the legal landscape. Suits for injuries from chemical exposure, for example, may require evidence on exposure effects from scientists with expertise in chemistry, biology, epidemiology, and pathology; a bridge collapse or a patent dispute may require engineering and technological expertise; and DNA evidence is often key in identifying criminals and excluding innocent individuals from prosecution. Moreover, science does not stand still. New developments in genetics, neuroscience, material sciences, and other fields are entering into legal discourse, and claims and cases are beginning to turn on them. As science has become, if anything, more important to the fair resolution of legal disputes, the quality of scientific evidence in the courts continues to be the subject of controversy.

In 1993, the U.S. Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* highlighted the obligation of judges to act as gatekeepers responsible for keeping unreliable scientific evidence from being admitted in litigation.[3] Following the *Daubert* decision, Judge Alex Kozinski, on remand, characterized the challenge for judges called upon to rule on the admissibility of expert scientific testimony:

> [T]hough we are largely untrained in science and certainly no match for any of the witnesses whose testimony we are reviewing, it is our responsibility to determine whether those experts' proposed testimony amount to "scientific knowledge," constitutes "good science," and was "derived by the scientific method."[4]

As Judge Kozinski's comments suggest and Justice Breyer's later observations indicate, *Daubert*, although it put more gatekeeping power in the hands of the judge, has far from resolved the tensions that arise when science appears relevant to litigation.[5]

Scientific societies have also focused on the stresses that exist between science and the law, often through the lens of ethics.[6] The American Psychological Association's code of conduct, for example, specifically addresses issues that arise when psychologists are called on to serve in forensic capacities.[7] The various, largely prosecution-oriented forensic sciences, spurred on by a critical National Academy of Sciences (NAS) report, have been working not only to increase the quality of their sciences but also to improve the accuracy and clarity of how forensic experts present their findings in court.[8]

A common explanation for complaints about the quality of the scientific evidence courts receive is the claim that "scientists tend to be leery of lawyers and the legal process, preferring not to venture into the courtroom."[9] Prior studies of experts in the American legal system provide some evidence of a disconnect between science and law, but the literature is sparse, consisting primarily of small surveys of testifying experts,[10] and four important case studies, each discussing cases from the pre-*Daubert* era: one involving an examination of court documents and interviews with the participants in six criminal and three civil cases that included scientific evidence,[11] and the other three analyzing court opinions in several cases involving statistical evidence.[12] Our current survey was designed to examine evidence for some of the themes touched on in this prior research (for example, dissatisfaction with the quality of opposing experts and questions about judicial competence) and to go beyond the prior research in examining in greater detail the response of experts to the legal system.

We designed our survey, in conjunction with the American Academy of Arts and Sciences' Public Face of Science project, to capture the views of distinguished scientists and engineers about the legal system and their experience with it. We surveyed scientists (including physical, biological, and social scientists) and engineers who were elected Fellows of the Academy.[13] We asked them whether lawyers or judges had ever requested their advice, whether they had ever agreed to help if asked, why they were willing to help and why they refused to provide help if they declined, and what their experience was if they assisted, and we sought their views on various aspects of the legal system and the system as a whole. We also explored their future willingness to participate in the legal system, and asked them whether certain proposed changes in legal procedures would affect that willingness to participate. Finally, we sought to determine whether participation correlated with and perhaps affected views of the legal system.

We were particularly interested in understanding how the legal system interacts (or doesn't) with the nation's most respected scientists and engineers. Not only are these people likely to have the most to offer the legal system, but if they are seen as willing to engage with the legal system, younger scientists and engineers may be more likely to follow. To capture the views of highly respected scientists

and engineers, we invited the members of the Academy in Class I (mathematical and physical sciences); Class II (biological sciences); and Class III (social sciences) to complete an online survey (n = 3328).[14] We obtained 366 responses, a response rate of 11.0 percent. The response rate is not as high as we had hoped, but our data constitute what is by far the largest number of scientists and engineers ever surveyed on their experience with, and perceptions of, the legal system.

Our response rate is similar to the 12.1 percent response rate that was obtained in a recent survey that sought to learn what members of another organization of scientists, the American Association for the Advancement of Science, thought about the FBI and law enforcement.[15] Hence we do not think the survey topic discouraged participation. To check for biases in our responding sample, we conducted a follow-up survey that could be answered in under five minutes, either by responding directly to questions on the email request or by going to a hyperlinked location like the one in the original survey. Two hundred fifty-three Academy members who had not responded to the original survey provided answers to this follow-up request. Those in our follow-up sample were similar to our sample respondents in gender, age, Academy class, whether they had ever been asked for assistance by the legal system, and how favorably they viewed the legal system. These similarities suggest that the experience and views of those who completed the initial full survey were not idiosyncratic. (See the methodological appendix posted at http://www.amacad.org/daedalus/whenlawcalls.) Moreover, this follow-up group gave us a larger total sample (n = 619) and a total response rate of 18.6 percent on which to examine participation rates and respondents' overall evaluations of the ability of the legal system to deal with science.

We also looked at how representative our respondents were by comparing the gender, age, and Academy class distributions of all Academy members and the initial sample. The distributions in the population and sample were substantially similar in these three categories. Sample respondents included a somewhat higher proportion of women (24 percent versus 17 percent).[16] And although the mean age in both the sample and population was seventy-one, the sample included a higher proportion of persons sixty-five or older (77 percent versus 69 percent) than is found in the overall population of Academy members.[17] The overrepresentation of those over sixty-five in the sample may reflect the less busy lives of partially or fully retired scientists, as well as the possibility that those who have in the past participated or been asked to participate as experts were more likely to respond than those without such experience, with older scientists likely having accumulated more opportunities to participate. Also, Class III members (social scientists and attorneys) responded at a somewhat higher rate than their proportion in the population (33 percent of respondents versus 28 percent of the population).[18] To see if these modest differences between the sample and population might distort our results, we conducted all analyses using both the unweighted responses and the responses weighted for gender, age, and class membership. Weighting did not change our results, so we use the unweighted data in presenting our findings.

While we cannot be certain that our sample respondents look like those Academy members who did not respond, there is little reason to suspect that the responses we received have serious relevant biases. Moreover, even if unknown biases exist, our survey sheds light on how a good proportion of the country's most distinguished scientists regard and interact with the legal system.

A majority (54 percent) of our respondents reported that they had been asked to provide expert scientific or engineering advice at least once. More than one-third (38 percent) said they had been asked three or more times, and one in six (17 percent) reported receiving ten or more requests.[19] If our nonrespondents were, as we expect, disproportionally people who were never asked for assistance, these rates are inflated; but note that a majority (60 percent) of respondents to our brief follow-up survey also said they had been asked for assistance. The request numbers suggest that the legal system approaches distinguished scientists and engineers for assistance with some frequency. Across disciplines, the most frequently asked experts worked in economics (87 percent), chemistry (81 percent), and engineering, computer sciences, and information technologies (80 percent). Next were noneconomist social scientists (72 percent). Those who reported the fewest requests were in the Academy's astronomy, physics, and earth sciences cluster (18 percent). Table 1 shows the full breakdown by disciplinary cluster. [20] These patterns make sense: experts in disciplines like astronomy are less likely to have expertise relevant to legal matters than experts in economics and chemistry.

When top experts are approached for assistance, they are likely to agree to provide it, at least on some occasions. In our sample, over 90 percent of those asked for advice agreed to assist at least once.[21] That willingness to serve is reflected in respondents' general agreement with the statement: "Absent strong reasons to the contrary, scientists should share their knowledge with the legal system when they are asked to serve as experts" (84 percent agreed or strongly agreed).[22]

About 10 percent of those who responded to our main survey never agreed to assist lawyers or judges when asked, while those who agreed to assist on one or more occasions may still turn down other requests. Why do they refuse? We asked respondents to check up to three of thirteen possible reasons for turning down requests, or to identify other reasons for refusing (Table 2). The most common reason for refusing to participate was "timing/other commitments" (66 percent). The demands faced by experts in legal matters can not only be time-consuming, but timing can also be unpredictable. Unlike experts who are full-time consultants or who are employed by the government to provide forensic expertise, professional scientists and engineers in both the academy and industry typically have jobs that make them only sporadically available to assist on legal issues. Strikingly few respondents mentioned formal organizational barriers to participation or advice against participating (6 percent), so it appears that few distinguished scientists are required by their employers' policies to turn down requests for assistance. Thus, it is time constraints rather than organizational restrictions that create a catch-22 for the legal system: the highest quality scientists have so much on their plates that they may be the least available to assist, even if they would otherwise be willing to do so.

The second most common reason for refusing to participate was that the "request was outside my area of expertise" (49 percent), an appropriate and desirable response since fit matters. The frequency of this response suggests that a system that helps lawyers and judges identify leading experts with knowledge specifically relevant to the issues in a case would increase the efficiency of searches for advice and might promote better expert advice in the legal system. In this connection, we asked those respondents who had provided assistance how, to the best of their knowledge, they had been identified by an attorney or judge as a potential expert.[23] Although commercial organizations provide directories of potential experts in various scientific and

Table 1
Academy Scientists Asked for Scientific or Engineering Advice, Requests by Discipline
Q: What is your field of scientific or engineering expertise?
Q: Has a party, attorney, or judge ever asked for your expert scientific or engineering advice?

	Ever asked for advice		
Fields of expertise	**Yes % (N)**	**No % (N)**	**Total % (N)**
Biological and cognitive sciences	50.5% (46)	49.5% (45)	100% (91)
Medical sciences	61.1% (11)	38.9% (7)	100% (18)
Astronomy, physics, and earth sciences	17.8% (8)	82.2% (37)	100% (45)
Chemistry	81.0% (17)	19.0% (4)	100% (21)
Mathematics and statistics	36.0% (9)	64.0% (16)	100% (25)
Social sciences except economics	71.8% (28)	28.2% (11)	100% (39)
Economics	86.7% (13)	13.3% (2)	100% (15)
Social and developmental psychology and education	57.1% (12)	42.9% (9)	100% (21)
Engineering, computer sciences, and information technologies	80.0% (20)	20.0% (5)	100% (25)
Law, including the practice of law	35.0% (7)	65.0% (13)	100% (20)
Total	53.4% (171)	46.6% (149)	100% (320)

engineering fields, attorneys, at least according to the respondents, rarely (6 percent) located them by using commercial referral sources. More commonly, respondents said they were identified through their scholarship, or their names were provided by another lawyer, another expert, or the client. It is likely, however, that scientists who are less publicly visible than Academy members and those for whom consulting is their primary professional activity would be more likely to be identified through commercial sources.

The next most common reason for refusal, offered by nearly one in four experts (24 percent) was that they "did not think the scientific or engineering evidence favored the party who wanted my knowledge." This response is inconsistent with willingness to be a "hired gun," a charge frequently leveled at expert witnesses. It may reflect the high quality of Academy experts and the fact that they do not need to rely on consulting for a dominant portion of their income. Expert refusals for this reason may have the positive consequence of leading attorneys to reassess the strength of their cases. They may, however, also encourage attorneys to search for more party-friendly

Table 2
Reasons for Turning Down Requests

Q: Thinking back to all the times you turned down requests to serve as an expert, what were your most common reasons for refusing? (Check up to three)

Reason	N Checked	% Checked
Timing/other commitments	89	65.9
Outside my area of expertise	66	48.9
Evidence didn't favor party asking	32	23.7
Doubts about the legal system (three items)	31	23.0
Particular parties or attorneys (two items)	28	20.7
Wanted my reputation, not my knowledge	28	20.7
Conflict of interest	15	11.1
Fee issues	10	7.4
Advice or institutional policy against (two items)	8	5.9
Other reasons	6	4.4
Total respondents (respondents could check up to three responses)	135	

experts, whether or not the party-friendly view has adequate scientific justification. Such searches, which can distort the quality or implications of the scientific evidence that finds its way into legal proceedings, are abetted by the absence of rules requiring attorneys to reveal the identities of all experts consulted in connection with a case. *Daubert* and its progeny should theoretically filter out the worst abuses of this sort, but the *Daubert* line of cases indicates it is a far from perfect filter.

Time constraints and mismatches are not the only reasons why the legal system loses potentially valuable scientific expert knowledge. Some experts indicated that they refused to assist because they had doubts about the legal system (23 percent). They questioned the ability of the adversary process to resolve science or engineering disputes, doubted whether the legal system could fairly resolve the dispute, or did not relish the prospect of being cross-examined. The majority (84 percent) of the respondents who expressed unease with the legal system had, however, agreed to assist in response to some requests and 68 percent had actually provided assistance. In some cases, their doubts were most likely stoked by their experiences.

Experts also turned down requests because they did not wish to assist particular parties or attorneys (21 percent). One respondent, for example, noted, "I will nev-

er work for a patent troll." To the extent that experts share preferences, some parties may find it difficult to obtain expert assistance.[24] An unequal supply of expertise may not undermine the quality of legal decision-making if expert preferences align with scientific merit, but it creates problems if they do not.

Respondents rarely identified fee issues as a reason why they refused requests for assistance (7 percent), although social desirability bias may have discouraged checking this response. It is, however, likely that fees are seldom the deal breaker for these scientists. As responses to this item indicate, other considerations seem to be more important. Not only are distinguished scientists and engineers likely to be able to command substantial compensation, at least in civil cases, but money may not be the principal motivator for the most successful, and typically the most highly paid, academic and industry scientists. Indeed, two respondents who cited fee issues said they refused to participate because "mostly attorneys did not want me to testify unless I would be paid, and I refused" and "[I] do not do this for the fees ever, but pro bono for the common good. Many requests I decline are for a fee which I do not feel appropriate to take." However, as we discuss below, promised financial compensation is a factor affecting the participation of some experts.

Taken as a whole, responses to our inquiry into why scientists choose not to participate in the legal system present a reassuring picture. Fewer than one in four of those refusing said they did so because of doubts about various aspects of the legal system, and only one respondent gave this as the sole reason for refusing to participate. Most often, the time needed to participate was a major factor (66 percent), and thirteen respondents (10 percent) gave time or organizational policies against participation as their only reasons for refusing. Perhaps most heartening is the degree to which ethical reasons appear to have motivated nonparticipation. These included admitted lack of expertise, feeling that the evidence did not favor the side that sought assistance, conflicts of interest, realizing that the lawyer making the request more highly valued the expert's reputation than knowledge, and not wanting to work for a particular client or attorney. Overall, 79 percent of our respondents listed at least one of these concerns as a reason for nonparticipation. There is almost no evidence in these data that the kinds of scientists elected to the Academy see themselves as, or are willing to be, "hired guns."

Participation as a testifying expert often involves a dramatic diversion from the central professional activities of Academy scientists and may be the most demanding role a scientific expert is called upon to play in the legal system. Our sample included ninety-four experts who indicated that in their most recent experience serving as an expert witness, they had testified in a hearing or trial. We asked them to evaluate the importance of various possible reasons for their willingness to participate as an expert in that case (see Table 3).

Consistent with a focus on scientific accuracy, the reasons our respondents rated as most important were the ability to assist in correctly resolving the case (85 percent) and the associated belief that the expert was testifying for the side that was scientifically correct (86 percent). Their side's moral correctness was an important reason for 72 percent of respondents, and more than half of respondents identified the obligation to share knowledge as an important motivation (64 percent). Nearly half (46 percent) said it was important that they thought it would be a learning experience. Only 30 percent said that wanting to affect law or policy was an important motivator.

A substantial minority (38 percent) said they viewed promised financial compensa-

Table 3
Importance of Reasons for Participating in Most Recent Case

Q: How important were each of the following reasons for your decisions to provide assistance in this case? (from 1 = Not very important to 5 = Extremely important)

Reasons for Participating (asked of those who indicated they had testified)	Responded with Important or Very Important
My expertise could assist in a correct resolution	85% (68/80)
Scientists have an obligation to share knowledge	64% (47/74)
My side was scientifically correct	86% (64/74)
My side was morally correct	72% (50/69)
A learning experience	46% (33/72)
Wanted to affect law or policy	30% (20/66)
Promised financial compensation	38% (25/66)

tion as an important factor motivating participation. Thus, although we found that experts seldom turned down requests to assist because they regarded the fees they would receive as insufficient, expert fees can be an incentive to participate. One expert explaining his participation commented, "I believe in sharing scientific knowledge and making legal decisions based on scientific knowledge, the cases are interesting, and I like the money." Another said, "I've been doing it for 40 years and overall greatly benefit from the experience. It enhances my research, teaching[,] collections of interesting life experiences, sense of helping the innocent and bank account." Several others said that in deciding whether to participate they considered both the time required and the level of compensation, with some noting they did not accept assignments when they felt their time would not be fairly compensated. Still others were quite blunt in describing the motivational effects of fees, including respondents who explained their willingness to participate in the future by writing, "compensation," "pay," and "[i]nterest, money." Still, when asked about the most recent case in which they testified, only 38 percent rated financial compensation as an important motivating reason, and most rated at least three other reasons as also important. Only one respondent gave money as the sole important motivation for providing assistance. Thus, although a few scientists refuse compensation for providing assistance, most expect to be compensated and many acknowledge that compensation is a motivator. Nonetheless, their motivations to assist do not appear to be driven solely or in most cases even largely by a profit motive.

Attorneys may ask the expert scientists and engineers they hire to testify at hearings or trials, to answer questions at depositions, or to write reports or affidavits. Reactions to these activities highlight differences between the legal system's demands and the way scientists and engineers generally spend their professional time.

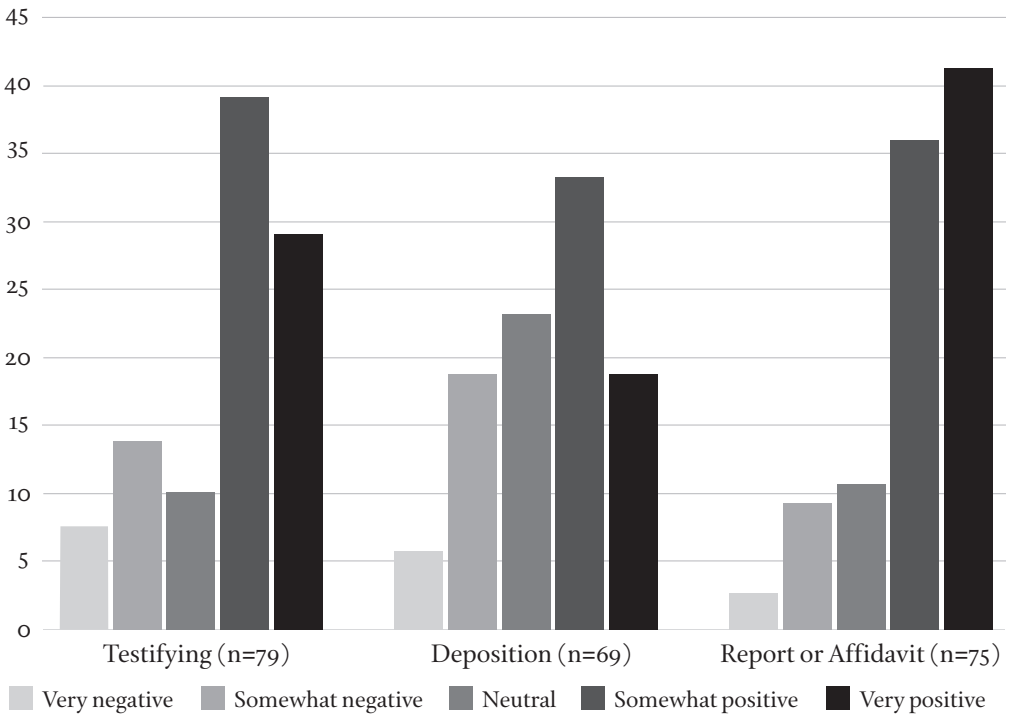
Court testimony typically follows a question-and-answer format. Unlike the classroom, where students may ask questions but the professor controls the flow (and number) of student remarks, in a courtroom the attorneys' questions seek to control how experts present their evidence and opinions. During direct examination, the questions come from the lawyer who hired the expert. This dialogue has typically been rehearsed, often incorporates the expert's suggestions, and is designed to persuade the judge and/or jury. In contrast, the opposing attorney's cross-examination typically attempts to constrain what the expert can say, sometimes in ways that will frustrate an expert whose strongest desire is to state the *whole* truth. The cross-examiner may also challenge not only the accuracy of the expert's opinions but sometimes the expert's competence and integrity as well. Not surprisingly, experts think more favorably of direct than cross-examination (81 percent versus 40 percent positive), and they see the lawyer on the side for whom they testified in a more positive light than the opposing side's lawyer (92 percent versus 31 percent positive).[25] Although some trial experiences generated complaints ("The entire process is reminiscent of a high school boy's locker room where attorneys try to play gottcha and to undermine rather than to reveal, reconcile, and allow the judge or jury to make informed decisions"), 68 percent of the experts who testified at trial rated the overall experience positively, including 29 percent who rated it very positively ("I enjoyed it – learned a lot – a different world").

In a deposition, unlike in a trial, only the opposing attorney asks questions, and no judge is present. Moreover, the rules of evidence, including rules of relevance, are relaxed. The expert in a deposition thus lacks the opportunity that a trial presents to educate a neutral decision maker, and is subject to cross-examination without a judicial referee to limit the nature or extent of the questioning. As with the trial experience, experts rated the deposition behavior of the attorney for their side more positively than the behavior of the opposing attorney (78 percent versus 25 percent). Some respondents, given the opportunity to elaborate, showed impatience with the experience ("only fishing expeditions"; "I don't like having my integrity questioned"). Experts who were deposed were on average less positive about the overall experience than those who testified: 52 percent rated it as positive, including 19 percent who rated it as extremely positive.

Unlike the trial testimony and deposition experience, report writing is familiar territory for scholars. Although an expert report or affidavit in litigation differs in form from that of a scholarly article, the expert in both instances is describing what she believes and the evidence supporting that belief. Experts by and large approved of (81 percent positive) the cooperation they received from the attorney who asked them to write a report. They reported that the attorney was willing to accept their independent view (91 percent) and that their report influenced the attorney's beliefs about the case (83 percent). Overall, writing a report or affidavit was the part of the litigation assistance process viewed most favorably (77 percent positive, including 41 percent very positive). Figure 1 shows how reactions to these three kinds of involvement differ.[26]

What we see reflects the generally positive view that expert participants have of their experience, but it also echoes a distaste for adversary procedures that some

Figure 1
Expert Evaluations of the Overall Experience: Testifying, Deposition, and Report or Affidavit Writing



experts identified as a reason why they refused to participate on one or more occasions.

We noted earlier that 90 percent of experts who had been asked for assistance had agreed to assist at least once. We also saw that experts often turn down invitations to serve. What about future service? We asked all respondents, "If you are asked in the future to serve as an expert in litigation, how likely is it that you would agree to serve?" One-third of our respondents (34 percent) said they were likely or very likely to serve, and 39 percent said they were uncertain. The remaining 28 percent said they were unlikely or very unlikely. We asked the ninety-five respondents who said they were unlikely to serve to tell us why they would be unlikely to serve. Of the eighty-five individuals who responded to this follow-up question,[27] sixteen mentioned being too old or that they had retired, and twenty-two mentioned being too busy, but thirty – one-third of these respondents – mentioned some distasteful reaction to courtroom behavior ("Accurate communication is extremely difficult and generally not desired by either side"; "Litigation sucks") or the adversary system ("Don't like being cross-examined"; "Because my experience was that my scientific expertise was not at issue – I was (unfairly) accused of inconsistent behavior"; "The experience of being deposed was horrible") or the inconsistent demands of science and law ("I am uncomfortable now in the adversarial system in courts dealing with

scientific matters"; "Often have difficulty with how scientific facts are distorted in legal proceedings to project what is wanted rather than what is true"). Twelve of the negative responses came from those who reported experience in providing expert assistance, while eighteen came from respondents who had no experience, thus reflecting a combination of responses to prior experience and images of the legal system not based on personal experience that mitigated against participation.[28]

Although many of these objections and sources of discomfort arise from intrinsic features of the American legal system and some are the legacy of a past unpleasant experience, other perceived problems may be open to adjustment. Thus, we assessed our respondents' reactions to potential changes in trial procedure that might make participation more attractive to experts. This effort focused on four procedural variations that might affect a respondent's willingness to participate in a legal proceeding.[29]

Being asked by a judge to serve as a court-appointed expert (see Daniel Rubinfeld and Joe Cecil's contribution to this volume) had the most appeal, leading more than two-thirds of the respondents (69 percent) to say that they would be more likely to serve if asked to be a court-appointed expert (Table 4).[30] This was particularly true among those who expressed uncertainty about future participation; 77 percent of those respondents said they would be more likely to participate if asked to serve the court rather than a party.[31] Moreover, few respondents, whatever their current inclination to serve, said they would be less likely to assist if the request came from a judge (2 percent overall).

A majority of respondents (59 percent) were also attracted by the idea of meeting privately with opposing experts and writing a joint report that indicated areas of agreement and disagreement. This option was particularly attractive to scientists currently uncertain about their future willingness to serve, leading 72 percent of them to say the change would make them more likely to participate.[32] Nonetheless, for some respondents, this change would decrease their willingness to serve (9 percent overall).

These two favored procedural modifications appear likely to diminish the adversarial nature of the expert experience. Court-appointed experts do not have partisan clients, and the opportunity to produce a joint report with the opposing expert potentially avoids or reduces clashes of expertise. The lesser enthusiasm for the third suggested change, permitting opposing experts to question one another in open court, is telling. Overall, less than one-third (32 percent) said it would increase their willingness to serve, and for one in five (22 percent), the change would make them less likely to serve. Even 14 percent of those who identified themselves as currently likely to participate said this procedural modification would make them less likely to serve. Thus, respondents expressed little interest in engaging in attorney-like adversary procedures by questioning and being questioned by an opposing expert. This is not because they reject all questioning. A majority of respondents (58 percent) liked the idea of allowing jurors to pose questions to them and few (3 percent) rejected it, perhaps because the procedure emulates a professor's availability to answer student questions. Overall, our results suggest that the supply of high-quality expertise can be expanded if the legal system creates procedural options that emulate scientific and academic exchange. Such procedural adjustments would reduce attorney control and may seem inconsistent with the traditional adversary system of the United States, but other common law countries with adversary systems, like Canada and Australia, have taken steps in this direction.[33]

Table 4
How Potential Procedural Modifications would Affect Future Willingness to Participate in Light of Current Willingness to Participate

	Current Willingness to Participate in the Future			
Change in Future Willingness to Participate in Response to Potential Procedural Modifications	**Unlikely to Participate in Future**	**Uncertain about Future Participation**	**Likely to Participate in Future**	**Total**
	27.8% (n=95)	**38.6% (n=132)**	**33.6% (n=115)**	**100.0% (n=342)**
If I were asked by a judge or arbitrator to serve as a court-appointed expert rather than by a party as an adversary expert:				
Would be more likely	63.6%	77.2%	63.2%	68.7%
No effect	34.1%	22.8%	32.5%	29.2%
Would be less likely	2.3%	0.0%	4.4%	2.1%
	100.0%	100.0%	100.1%	100.0% (n=329)
If I were permitted to meet privately with opposing experts to discuss issues and write a joint report indicating areas of agreement and areas of disagreement:				
Would be more likely	45.5%	72.2%	55.0%	59.1%
No effect	46.6%	19.8%	33.3%	31.7%
Would be less likely	8.0%	7.9%	11.7%	9.2%
	100.1%	99.9%	100.0%	100.0% (n=325)
If I could question opposing experts in court and they could question me:				
Would be more likely	25.3%	33.1%	37.2%	32.4%
No effect	50.6%	40.3%	48.7%	46.0%
Would be less likely	24.2%	26.6%	14.2%	21.6%
	100.1%	100.0%	100.1%	100.0% (n=324)
If I could answer juror questions after I gave my testimony:				
Would be more likely	44.3%	63.2%	62.2%	57.7%
No effect	50.0%	34.4%	36.9%	39.5%
Would be less likely	5.7%	2.4%	0.9%	2.8%
	100.0%	100.0%	100.0%	100.0% (n=324)

We have seen that the scientists in this survey often expressed frustration with legal procedures and, in some cases, a suspicion that those procedures were purposefully designed to avoid getting at the truth. How did the scientist-respondents as a whole view the success of the legal system in producing decisions that accord with sound science? Overall, we found that 60 percent of our respondents saw the legal system as very or somewhat successful while 40 percent had the opposite view.[34] What explains this division of opinion? One possibility is that experience with the legal system leads to greater familiarity and more positive attitudes. Another is that experience and familiarity engender disappointment and cynicism, evoking more negative attitudes. As a first step, we compared the attitudes of those with and without experience providing advice. Those with experience rated the legal system as significantly more successful, with 70.0 percent of participants seeing the system as somewhat or very successful, while only 53.5 percent of the nonparticipants expressed that favorable view.[35] This difference was also reflected in other perceptions and attitudes toward the legal system. Participants rated lawyer understanding of science more favorably than nonparticipants, saw scientists as treated with more respect, and viewed serving as an expert witness more favorably as a way to keep abreast of the real world implications of their science. Participants did, however, express somewhat greater criticism for experts, indicating greater agreement than nonparticipants with the belief that even respected experts may compromise their standards in the context of the legal system (Table 5).

Although this overall pattern undercuts the hypothesis that experience tends to undermine confidence in the legal system, we cannot be certain that it promotes it. People may agree to participate because they view the legal system positively (selection effect), their view may be shaped by their participation (experience effect), or both may help explain the correlation.

A modest quasi control group bears on the relative plausibility of the selection and experience effects (Table 6). Thirty-two respondents agreed at least once to participate but never actually participated. We did not ask why their agreement did not result in participation, but given how the litigation process works, we expect the most common reason is that the case was withdrawn or there was a quick settlement or plea agreement. The pattern of responses from this agreed-but-never-participated group was closer to the never-asked group than to the group of participating respondents (Table 6).

The groups differed significantly on four statements in Table 6 (different subscripts indicate significant differences on the post hoc comparisons). In each of these comparisons, the "never asked" and "participated" groups differed from one another. On the evaluation of lawyer understanding, the participated group was distinctive: only participation was associated with an increased evaluation of the ability of lawyers to understand science. This pattern is consistent with an increased appreciation of how well lawyers understand science arising from close interaction. It may also be a biased view of how well lawyers understand science since those lawyers who hired scientific experts and worked with them may be better able to grasp scientific concepts than the general run of attorneys.

Most important, we compared the groups on their views about the success of the legal system in dealing with scientific matters. Again, the participants viewed the legal system as more successful (70.0 percent) than both those never asked (52.5 percent) and those who agreed but did not have an opportunity to participate (51.6 percent). The pattern is only suggestive in light of the small number of quasi control respondents

Table 5
Perceptions of and Attitudes toward the Legal System by Participants and Nonparticipants*

	Never Participated (n=201)	Participated (n=124)	p-level**
Science should aid law[a]	4.09	4.00	ns
Judges can understand science[b]	2.81	2.85	ns
Jurors can understand science	2.44	2.39	ns
Lawyers can understand science	2.80	3.18	$p < .001$
Scientists are treated with respect[c]	3.14	3.43	$p < .002$
Experts compromise standards[d]	3.17	3.37	$p < .05$
Links real world and science[e]	2.75	3.12	$p < .003$
Success of legal system with science (% successful)[f]	53.5%	70.0%	$p < .002$

* Scale ranges from 1 (strongly disagree) to 5 (strongly agree)

** "p <" = significant level; ns = not significant at the .05 level

[a] "Absent strong reasons to the contrary, scientists should share their knowledge with the legal system when they are asked to serve as experts."

[b] "In cases where science is important to the decision, most judges and arbitrators have the ability to understand scientific evidence and the scientific process." The next two items substitute "most juries contain jurors who" and "most lawyers" for "most judges and arbitrators."

[c] "Scientists are treated with appropriate respect when they testify at trials or in depositions."

[d] "Even respected scientific and engineering experts may compromise their scientific standards and write reports or give testimony [that] better support the position of the party that hired them."

[e] "Serving as an expert witness is a good way for scientists to keep abreast of the real world implications of their sciences."

[f] "In litigation or arbitration where scientific or engineering issues are involved, on average, how successful do you think the American legal system is in producing results that reflect sound scientific or engineering knowledge?" (percent somewhat or very successful).

Table 6
Perceptions of and Attitudes toward the Legal System with Quasi Control Group*

	Never asked (n=152)	Participated (n=124)	Asked and agreed but did not participate (n=32)	Overall p-level**
Science should aid law	4.14	4.00	3.97	ns
Judges can understand science	2.80	2.85	2.94	ns
Jurors can understand science	2.42	2.39	2.41	ns
Lawyers can understand science	2.82_a	3.18_b	2.75_a	$p < .005$
Scientists are treated with respect	3.15_a	3.43_b	3.24_{ab}	$p < .01$
Experts compromise standards	3.18	3.37	3.10	ns
Links real world and science	2.72_a	3.12_b	2.90_{ab}	$p < .007$
Success of legal system with science	$52.5\%_a$	$70.0\%_b$	$51.6\%_a$	$p < .01$

* Scale ranges from 1 (strongly disagree) to 5 (strongly agree)

** "p <" = significant level; ns = not significant at the .05 level

Note: Subscripts indicate significant differences on the post hoc comparisons.

and the unknown reasons why they did not end up participating. Nevertheless, we provided an opportunity to support the possibility that our results were the result of preexisting views of the legal system, and the data fell in the opposite direction.

This survey provides unique information about how scientists interact with and view the legal system. There are aspects of our data that we have yet to plumb, but even after further analysis, we must be careful in generalizing from our results: The findings we report may characterize only, or largely, the kinds of scientists who achieve substantial success in their fields. We do not know how scientists who market themselves as scientific experts, including scientists who work for consulting firms or the large group of forensic scientists who testify regularly for the prosecution, would answer the questions we posed.[36] Also, given the age and accomplishments of Academy members who are scientists, we cannot be certain how the generation of scientists now entering the most productive portions of their careers view the legal system or would respond to proposed changes in legal procedure. Nevertheless, the snapshot we provide of the group of eminent scientists who responded

to our survey is an important one. Our respondents have expertise that is crucial for a legal system that must increasingly take account of scientific understandings and will be well served only if the science available to it is both clear and sound.

In this respect, the good news is that the Academy survey reveals that the legal system has often been able to draw on distinguished scientists and engineers for assistance when scientific and engineering questions intersect with the law. This capacity can be expected to continue into the future. When asked, most scientific experts are willing to participate in legal actions, at least some of the time. Still, the relationship has its trouble spots, including some discomfort with the adversary system, that seem to reflect the different cultural norms of science and law. Although our survey responses suggest that several modest changes in trial procedures could have positive effects for both experts and triers of fact, as other essays in this volume indicate, tensions between science and the law are unlikely to ever completely disappear.

AUTHOR BIOGRAPHIES

SHARI SEIDMAN DIAMOND, a Fellow of the American Academy since 2012, is the Howard J. Trienens Professor of Law and Professor of Psychology at the Northwestern University Pritzker School of Law, and Research Professor at the American Bar Foundation. She is the author of the *Reference Guide on Survey Research* in the *Reference Manual on Scientific Evidence* (3rd ed., 2011) and *The Multiple Dimensions of Trial by Jury: Studies of Jury Behavior* (2016, in Spanish) and editor of *Trademark and Deceptive Advertising Surveys: Law, Science, and Design* (with Jerre B. Swann, 2012).

RICHARD O. LEMPERT, a Fellow of the American Academy since 1993, is the Eric Stein Distinguished University Professor of Law and Sociology, emeritus, at the University of Michigan. He is the author of *A Modern Approach to Evidence: Text, Problems, Transcripts, and Cases* (with Samuel Gross et al., 5th ed., 2013) and *An Invitation to Law and Social Science* (with Joe Sanders, 1986) and editor of *Evidence Stories* (2006).

AUTHORS' NOTE

Our thanks to Beth Murphy, American Bar Foundation, and Eleanor Wilking, Northwestern University Pritzker School of Law, for their excellent assistance. We are grateful to the American Academy of Arts and Sciences and to the American Bar Foundation for funding support. We are also grateful for valuable feedback from the authors of the other essays in this volume who were participants in the American Academy of Arts and Sciences meeting.

ENDNOTES

[1] *The tryal of Spencer Cowper, Esq, John Marson, Ellis Stevens, and William Rogers, gent. upon an indictment for the murther of Mrs. Sarah Stout, a Quaker before Mr. Baron Hatsell, at Hertford assizes, July 18, 1699*, 45, http://quod.lib.umich.edu/e/eebo/A63196.0001.001/1:3?rgn=div1;view=fulltext. The judge went further: "Dr. Brown has a learned discourse in his Vulgar Errors upon this subject, concerning the floating of dead bodies, I don't understand it my self, but he hath a whole chapter about it." Ibid., 16.

[2] Stephen Breyer, "The Interdependence of Science and Law," *Science* 280 (5363) (1998): 537.

[3] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

[4] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 43 F.3d 1311, 1316 (9th Cir. 1995), *cert. denied*, 116 S.Ct. 189 (1995).

[5] Sophia I. Gatowski, Shirley A. Dobbin, James T. Richardson, et al., "Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-*Daubert* World," *Law & Human Behavior* 25 (5) (2001): 433; and John Meixner and Shari Seidman Diamond, "The Hidden *Daubert* Factor: How Judges Use Error Rates in Assessing Scientific Evidence," *Wisconsin Law Review* 2014 (6) (2014): 1063.

[6] Andre A. Moenssens, "Ethics: *Codes of Conduct for Expert Witnesses*," in *Wiley Encyclopedia of Forensic Science*, ed. Allan Jamieson and Andre A. Moenessens (Hoboken, N.J.: Wiley-Blackwell, 2016).

[7] American Psychological Association, "Ethical Principles of Psychologists and Code of Conduct" [amended February 20, 2010], *American Psychologist* 57 (2002): 1060–1073.

[8] National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (Washington, D.C.: National Academies Press, 2009). The federal government responded to the NAS report on the state of the forensic sciences first with the formation of an interagency task force to consider the report's critiques and recommendations, and then with the establishment of the National Commission on Forensic Sciences, a distinguished panel of academics, judges, prosecutors, forensic scientists, and defense counsel, charged with assessing and ensuring the scientific integrity of the various forensic sciences. The Commission was, however, terminated after President Trump took office. Still ongoing, as we write, is a National Institute of Standards and Technology–Department of Justice effort in which representatives of the various forensic sciences are working to establish performance and communication standards for their disciplines.

[9] National Research Council, *A Convergence of Science and Law* (Washington, D.C.: National Academies Press, 2001). The concern of courts and the legal academy is not that lawyers cannot find experts. Rather, the concern is with the pool of available experts, their biases and scientific competence, and their effectiveness in communicating their scientific opinions fairly and clearly in court. Hence the Council's concern is with factors that discourage the most able scientists from lending their expertise to the courts.

[10] A two-part survey: Anthony Champagne, Daniel Shuman, and Elizabeth Whitaker, "An Empirical Examination of the Use of Expert Witnesses in American Courts," *Jurimetrics Journal* 31 (4) (1991): 375; and Daniel W. Shuman, Elizabeth Whitaker, and Anthony Champagne, "An Empirical Examination of the Use of Expert Witnesses in the Courts – Part II: A Three City Study," *Jurimetrics Journal* 34 (2) (1994): 193. See also Jonathan Baker and M. Howard Morse, *Final Report of Economic Evidence Task Force*, Task Force on Economic Evidence (Chicago: American Bar Association, 2006), Appendix II.

[11] Michael J. Saks and Richard Van Duizend, *The Use of Scientific Evidence in Litigation* (Williamsburg, Va.: National Center for State Courts, 1983).

[12] Stephen E. Fienberg, ed., *The Evolving Role of Statistical Assessments as Evidence in the Courts* (New York: Springer, 1989). Joseph B. Kadane and Caroline Mitchell, "Statistics in Proof of Discrimination Cases" and Richard O. Lempert, "Befuddled Judges: Statistical Evidence in Title VII Cases" in *Legacies of the 1964 Civil Rights Act*, ed. Bernard Grofman (Charlottesville: University of Virginia Press, 2000), 241–262 and 263–281, respectively, both analyze court responses to expert evidence in the same seventeen cases.

[13] We developed the survey with assistance from John Randell and Keerthi Shetty at the American Academy. We revised our draft and improved it substantially following conference calls with scientists, engineers, and legal scholars who had reviewed the original draft of the survey, as well as with the benefit of a conversation with Justice Stephen Breyer (December 2015 – January 2016). Robert Townsend of the American Academy formatted and distributed the online survey in two waves (April 2016 and September 2016).

[14] This group also included sixty Academy members from Class V (public affairs, business, and administration) whose substantive expertise was in science or engineering.

[15] Nathaniel Hafer, Cheryl J. Vos, Karen McAllister, et al., "How Scientists View Law Enforcement," *Science Progress*, February 2009. (The American Association for the Advancement of Science and the American Academy of Arts and Sciences share the AAAS acronym, but are

different organizations with different memberships. The Association in the Hafer et al. survey is a voluntary subscription organization, while the Academy Fellows we surveyed were scientists nominated by their peers and voted into the Academy based on their professional accomplishments.) See also Brian J. Love, "Do University Patents Pay Off? Evidence from a Survey of University Inventors in Computer Science and Electrical Engineering," *Yale Journal of Law and Technology* 16 (2) (2014): 285, 299 [reporting an 11.3 percent response rate].

[16] Chi-squared = 9.16, $p < .003$. The higher response rate for women is not unusual: see, for example, William G. Smith, *Does Gender Influence Online Survey Participation? A Record-Linkage Analysis of University Faculty Online Survey Response Behavior* (San José, Calif.: San José State University, 2008), http://files.eric.ed.gov/fulltext/ED501717.pdf; it also occurred in the follow-up sample (23 percent women).

[17] Chi-squared = 8.14, $p < .005$. Although the mean age in the follow-up survey was seventy-one, the distribution of respondents was somewhat closer to the population, with 74 percent aged sixty-five or older.

[18] In contrast, Class III members were somewhat underrepresented in the follow-up survey (26 percent), so that together the two surveys had representation from Class III that was similar to the population (30 percent versus 28 percent). The other two classes had nearly identical representation in the population, first, and second surveys (Class I: 37.5 percent, 35.5 percent, 39.1 percent; Class II: 34.1 percent, 31.1 percent, 34.8 percent).

[19] Sixty percent in the follow-up survey reported they had been asked at least once, 33 percent at least three times, and 12 percent ten or more times.

[20] These percentages are based on disciplinary categories with at least ten respondents. We did not obtain disciplinary information in the follow-up survey.

[21] In the follow-up survey, 84 percent of respondents agreed to assist at least once. Accepted invitations did not always result in participation: 20 percent of those in the original sample and 4 percent in the follow-up survey who agreed did not end up participating.

[22] Respondents could also answer that they were undecided, disagreed, or strongly disagreed. Agreement with some other statements was considerably less; for instance, agreement rates with assertions that most judges and arbitrators (30 percent), jurors (12 percent), and lawyers (39 percent) have the ability to understand scientific evidence and the scientific process.

[23] Respondents were asked to choose all applicable sources from the following list: my scholarship (77 percent); an expert referral organization (6 percent); a referral from another lawyer (21 percent); name provided by the client (23 percent); recommended by another expert (22 percent); don't know (7 percent).

[24] Joseph Sanders, Betty Rankin-Widgeon, Debra Kalmuss, and Mark Chesler, "The Relevance of 'Irrelevant' Testimony: Why Lawyers Use Social Science Experts in School Desegregation Cases," *Law and Society Review* 16 (3) (1981–1982): 403 [plaintiff lawyers appear to have easier access to a network of scholars willing to testify].

[25] Respondents rated various features of the trial, deposition, or report/affidavit on a 5-point scale: very negative, somewhat negative, neutral, somewhat positive, very positive.

[26] A subset of thirty-five respondents provided all three ratings because their most recent experience had required a report, a deposition, and testimony. The pattern for this subset mirrored the results in Figure 1.

[27] Two of those who said they were unlikely to serve were federal judges and eight who said they were unlikely to serve did not indicate why.

[28] The third and seventh quotes in the text came from respondents without experience; the remaining came from respondents with prior experience.

[29] For each potential change, respondents were asked: If [change was made], I would definitely be more likely to participate; I would probably be more likely to participate; It would have

no effect on my decision whether to participate; I would probably be less likely to participate; I would definitely be less likely to participate.

[30] Daniel L. Rubinfeld and Joe S. Cecil, "Scientists as Experts Serving the Court," *Dædalus* 147 (4) (Fall 2018).

[31] Chi-squared$_4$ = 10.51, p < .04.

[32] Chi-squared$_4$ = 19.74, p < .001.

[33] See Ian Freckelton QC, Jane Goodman-Delahunty, Jacqueline Horan, and Blake McKimmie, *Expert Evidence and Criminal Jury Trials* (Oxford: Oxford University Press, 2016), chap. 3.

[34] Respondents were asked to rate the legal system's success by choosing among four options: very successful, somewhat successful, somewhat unsuccessful, very unsuccessful.

[35] Chi-squared = 8.31, p < .004. This pattern was replicated in the follow-up survey. Sixty-eight percent of participants viewed the legal system as successful, while 61 percent of the nonparticipants did; Chi-squared for the combined samples = 9.15, p < .002. Neither age nor gender was associated with this view. A majority of all three Academy classes viewed the system as successful, although Class II members (biological sciences) were least positive (55.8 percent), Class I members (mathematical and physical sciences) were more positive, and Class III (social sciences) were most positive (67.9 percent) (Chi-squared$_2$ = 5.36, p < .07). Nonetheless, within each class, those who had participated as experts in the legal system were more likely to view the legal system as successful than those who had not. When age, gender, Academy class, and participation are used to predict judged success, only participation is a significant predictor (Wald = 7.09, p < .01).

[36] Among those respondents who said they had provided assistance, most (84 percent) said they had assisted primarily in civil cases, 11 percent primarily in criminal cases, and 5 percent in both about equally. Assistance in civil cases was fairly evenly divided, with 33 percent primarily assisting plaintiffs, 27 percent primarily assisting defendants, and 40 percent assisting both about equally. Among the small group of 14 respondents who reported experience assisting in criminal cases, half (7) primarily assisted the defense, 4 primarily the prosecution, and 3 both sides about equally. The Academy sample thus included little if any representation from the large cadre of government-employed forensic scientists who regularly appear in criminal court cases.

Law & Neuroscience: The Case of Solitary Confinement

Jules Lobel & Huda Akil

Abstract: This essay discusses the interface between neuroscience and the law. It underscores the potential for neuroscience to break down the division that currently exists in law between physiological and psychological harm and between physical and mental injury. To show how scientific knowledge can illuminate a complex legal issue, we analyze the recent use of neuroscience in evaluating the harm caused by prolonged solitary confinement.

JULES LOBEL holds the Bessie McKee Walthour Endowed Chair at the University of Pittsburgh School of Law. He is the author of *Less Safe, Less Free: The Failure of Preemption in the War on Terror* (with David Cole, 2007), *Success Without Victory: Lost Legal Battles and the Long Road to Justice in America* (2003), and *A Less than Perfect Union: Alternative Perspectives on the U.S. Constitution* (1988).

HUDA AKIL, a Fellow of the American Academy since 2004, is the Quarton Professor of Neurosciences and Codirector of the Molecular & Behavioral Neuroscience Institute at the University of Michigan. She has published articles in such journals as *Nature*, *Nature Medicine*, *Science*, and *Journal of Neuroscience*.

Neuroscience is increasingly used in the courtroom, in a variety of circumstances.[1] Over the past decade or so, the distinct field of "law and neuroscience" has developed (sometimes termed "neurolaw"), a casebook on law and neuroscience has been published, courses on the subject are being taught in law schools and other departments, and the John D. and Catherine T. MacArthur Foundation has invested over $15 million in developing the Law and Neuroscience Project and Research Network.[2] Neuroscience testimony in the courtroom has, to date, largely focused on issues relating to criminal responsibility, with defense attorneys seeking to introduce brain scans of defendants to show that either they were not responsible for their actions or to argue that brain defects or problems justified mitigated penalties.[3]

Possible uses of neuroscience in the law go far beyond criminal cases, however. Neuroscience has the potential to bridge the divide in American law and culture between physical and mental injuries. For instance, it could enable judges to allow plaintiffs to recover damages in tort actions where mental harm may be uncompensable or disbelieved, but provable brain damage can be viewed as a physical injury.[4] Brain damage can be structural, such as a tumor or dimin-

doi:10.1162/DAED_a_00520

ished volume of a particular brain region, and/or it can be functional, such as a characteristic change in the activity of a brain circuit implicated in certain conditions, including severe chronic stress or depression, chronic pain, or loss of cognitive function. So, too, neuroscience might be useful in helping judges to understand the mental harms that government action can inflict and to determine whether the infliction of mental harm, intended or not, rises to the level of a constitutional violation.

This is already happening in one area: expert neuroscience evidence is being mustered to support claims of extreme and long-lasting, if not permanent, mental harm in constitutional challenges to prolonged solitary confinement, a disciplinary practice used in many state and federal prisons. Thus, in the class action case of *Ashker v. Governor*, challenging the solitary confinement of more than one thousand prisoners at Pelican Bay State Prison in California, the plaintiffs submitted expert neuroscience testimony in support of their Eighth Amendment claims that such prolonged confinement constitutes cruel and unusual punishment.[5] This essay reviews the current intersection between the law and neuroscience and then explores and analyzes neuroscience's use in evaluating the harm caused by prolonged solitary confinement.

At first, the connection between the law and neuroscience may seem surprising; the "Law and neuroscience seem strange bedfellows."[6] As legal scholar David Faigman has noted, there is a "fundamental divide between the fields of neuroscience and law," an observation that could also be made about the law and other fields in mainstream science.[7] Neuroscientists study the brain and are generally unconcerned with legal questions, while lawyers, as smart as they may be, usually know nothing about how the brain works and are not troubled by their ignorance. Yet the law and lawyers are ultimately concerned with regulating human behavior, and issues of intent are part of the grist in the legal mill. Understanding the brain is central to both the law and neuroscience; thus, the burgeoning interplay between the two fields should not be surprising.

Perhaps the most salient source of tension between the two fields has to do with the differing goals of the scientist and the lawyer. The scientist studying the brain is ideally a neutral analyst, an empiricist who pursues evidence to generate a better understanding of brain function regardless of preconceptions. The lawyer is ordinarily not neutral, but rather is an advocate for his or her client's interests. A scientist is only supposed to draw a definitive conclusion when findings are replicable to a very high degree. Yet lawyers and judges are seldom in a position to withhold judgment. They can, and often must, evaluate evidence bearing on a claim, even if it is not conclusive. Moreover, in civil cases, the usual standard of proof is not the scientific standard, which demands substantial certainty, but rather the preponderance of evidence, which translates into "more likely than not."

This difference leads to tensions that bear on both the potential uses and the need for caution when using neuroscience evidence in legal contexts. Lawyers would like to present favorable neuroscience evidence as dispositive, yet scientific norms specify that neuroscience claims should not be oversold. This does not mean that the neuroscientist cannot or should not advocate positions based on the science as we know it now, even if current science provides only strongly probable but not scientifically conclusive confirmation of a relationship. It does, however, mean that the neuroscience expert must admit, and indeed should proactively bring forth, the existence of scientifically sound conflicting evidence or underscore areas where current knowledge is

either lacking or too weak to support strong conclusions. In these circumstances, neuroscience advocacy is most likely to be relied upon by courts when its conclusions are consistent with common sense.[8]

Neuroscientific evidence has been used with significant success to mitigate punishment, particularly in capital cases.[9] In the juvenile death penalty case *Roper v. Simmons*, the Supreme Court seems to have utilized such evidence in support of its decision that it is unconstitutional to impose capital punishment on a minor.[10] Yet some of the more radical claims made by neuroscientists, like the claim that brain imaging undermines the whole basis of criminal responsibility, have been deeply controversial and have not gained much traction in the courts.[11] Moreover, outside of the criminal mitigation context, most efforts to introduce neuroscience evidence in courts have proven unsuccessful.[12] Nonetheless, neuroscience evidence continues to be introduced in civil cases.

There appear to be two broad ways in which neuroscience evidence has made its way into the legal system. The first is the use of case-specific evidence from brain imaging, such as Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) scans, to demonstrate a particular criminal defendant's defective ability to make rational decisions or to show harm to the brain suffered by a plaintiff.[13] The second use, more important to this essay, is as what has been termed *framework* or *foundational scientific evidence*: scientific testimony bearing on how other evidence should be used based on general theories or hypotheses.[14]

These and other uses of neuroscientific evidence have the potential to break down the sharp dividing line the law has erected between mental injury and bodily harm. In diverse fields of law, from torts to constitutional law, the legal system treats mental harm differently from physical harm. Tort law traditionally makes a distinction between physical and emotional harm, "with emotional harm being treated as a second class citizen."[15] For example, to recover for the negligent infliction of emotional distress, a majority of states require the plaintiff to show not merely mental or emotional harm, but also physical injury.[16] The reasoning that the courts generally provide for this limitation is that mental harm, unlike physical injury, is essentially subjective and therefore the physical injury requirement will give "a sufficient basis for the trial courts to determine [that the claims of mental harm are] not . . . fraudulent claims."[17] Although often reaffirmed, this nexus requirement emerged many years ago, long before the capabilities that modern neuroscience gives us existed.

The American Law Institute's recent Third Restatement of Torts incorporates as a general rule this clear distinction between physical or bodily injury and mental or emotional injury.[18] It does, however, allow for claims of intentional or negligent infliction of pure, stand-alone emotional harm, but only in very circumscribed circumstances, citing, among other things, concerns that "emotional harm is less objectively verifiable than physical harm" and that "some degree of emotional harm is endemic to living in society."[19]

Neuroscience research at least muddies the distinction between bodily injury and mental harm, and, in the future, it might negate it entirely. One tool that neuroscience can deploy is brain imaging, which allows a window into the altered functioning of the brain under different conditions. This approach has been used to study chronic pain, considered the greatest source of disability worldwide. Neuroimaging has shown that chronic pain does indeed change brain function, altering specific neural pathways broadly, leading some to classify it as a neurodegenerative disorder. The brain changes resulting from chronic pain may not yet

reach the standard of being diagnostic on their own. Nevertheless, they are reliable enough to motivate recent reviews putting forward neuroimaging strategies as a potential basis of evidence for both clinical and legal purposes.[20] It is notable that emotional suffering, including chronic anxiety and depression, has an equally profound impact on brain structure and function. Indeed, some of the same brain regions are disrupted in both chronic pain and depression, providing clear biological evidence of the overlap between physical and mental distress.[21]

Other types of mental harm such as Post-Traumatic Stress Disorder (PTSD) can be shown objectively to affect the brain, thereby demonstrating that this emotional injury is also physical in nature. Indeed, one court has so ruled. In the Michigan case *Allen v. Bloomfield Hills School District*, the plaintiff was operating a train when he crashed into a bus that had negligently strayed onto the train tracks. The plaintiff developed PTSD because the crash resulted in the deaths of several schoolchildren. The lower court dismissed his tort claim because the applicable Michigan statute required a showing of "bodily injury," which the court ruled the plaintiff had not proved.[22]

The Court of Appeals reversed the ruling, relying on PET scans of the plaintiff, showing that he had suffered abnormalities in the brain due to the accident.[23] The court noted that "brain injury is a bodily injury."[24] The "plaintiff presented objective medical evidence that a mental or emotional trauma can indeed result in physical changes to the brain. . . . There should be no difference medically or legally between an objectively demonstrated brain injury, whether the medical diagnosis is a closed head injury, PTSD, [or] Alzheimer's Disease."[25] The brain is a part of the body, and hence an injury to the brain that is objectively verifiable should count as physical injury.

The neuroscientific insight that mental pain and harm are sometimes the result of or correlated with brain damage or abnormalities may also play an important role in constitutional jurisprudence addressing American prison systems' practices of prolonged solitary confinement.

At any given time, an estimated one hundred thousand prisoners in this country are held in solitary confinement. Such confinement varies slightly from state to state, but it generally involves a prisoner being kept for approximately twenty-three hours a day alone in a small cell, with minimal social contact and no physical contact with others.[26]

A draconian example of such solitary confinement existed for many years at the Pelican Bay State Prison Security Housing Unit (SHU). At that prison, built in 1989, approximately 1,300 prisoners were imprisoned in small, Spartan, eighty-square-foot cells with no windows for almost twenty-three hours a day. For years, they had no view of the outside world; they saw no birds, trees, cars, or grass.[27] For one-and-a-half hours per day, they went out to a recreation "yard" attached to their cell block. This was a facility about twice the size of their cell, with fifteen-foot-high walls and a grate over the top where they recreated, alone. If they went out to the yard at the right time during the day, it was possible to see a little sunlight, but, generally, most prisoners had only fleeting, if any, glimpses of direct sunlight during their stay at Pelican Bay. They were allowed no phone calls at all except in an "emergency," which was defined as a parent dying, in which case they were allowed a fifteen-minute call with next of kin. They were permitted visits with their family, but no contact visits, meaning they only could speak with their visitors through an intercom, viewing them through a glass window, unable to touch or hug their loved ones. While some had televisions and radi-

os, there was no educational, vocational, or religious programming or activities.[28]

One might think that only the most heinous, pathologically violent prisoners would be placed in these conditions. But, in fact, most of the 1,300 prisoners at the Pelican Bay SHU were not there because of any violent act they had committed in prison, but solely because they were either members or associates (a loose definition that included people who simply associated with members) of a prison gang. These prisoners were placed in the SHU for an indeterminate period of time, which in practice generally meant until the end of their prison terms, unless they were paroled, snitched, or died. In short, the only real way out of the SHU and into the general prison population was to become an informant against the gang, usually a dangerous proposition.

It is hard to imagine surviving in this environment for more than a few days or weeks without becoming suicidal or mentally ill. Some of the prisoners placed in the SHU did become mentally ill. But hundreds did not. It is a testament to the human being's ability to adapt to atrocious conditions that many prisoners were able to survive these conditions not only for weeks, but for decades. As of 2011, almost one hundred of the prisoners at Pelican Bay SHU had been held in solitary confinement for over two decades, and almost five hundred had been so confined for more than ten years. Survival does not, however, mean that they did not suffer serious mental harm: depression, paranoia, and loss of concentration and memory are just some of the symptoms associated with extended solitary confinement.

In 1990, within a year after the Pelican Bay SHU opened, a high-powered and skilled group of lawyers sued the California prison system on behalf of the class of prisoners incarcerated at the Pelican Bay SHU. They drew as the judge who would hear the case one of the most progressive, civil-rights oriented federal judges in the entire country, Thelton Henderson. The case went to trial in 1993, and in early 1995, Judge Henderson ruled that California officials had denied plaintiffs' constitutional rights by using excessive force and by not providing adequate medical care.[29] Yet on the fundamental issue of whether placing prisoners in such strict isolation for years by itself constituted cruel and unusual treatment prohibited by the Eighth Amendment, Henderson did not pull the trigger, even if he did find that the conditions were draconian, sterile, and isolating. For example, he opined that "the overall effect of the SHU is one of stark sterility and unremitting monotony."[30] He found that the conditions of social isolation were profound and noted that when he visited the prison, he observed prisoners pacing around in their cells as if they were animals in a zoo.[31]

The plaintiffs had submitted expert testimony from two internationally prominent psychological experts who had interviewed many prisoners and concluded that they suffered from varying degrees of psychological pain, including paranoia, lack of concentration, chronic depression, confused thought processes, hallucinations, irrational anger, emotional flatness, violent fantasies, and oversensitivity to stimuli.[32] Henderson acknowledged that mental pain, but held that it did not rise to the level of a constitutional violation, stating:

> the record demonstrates that the conditions of extreme social isolation and reduced environmental stimulation found in the Pelican Bay SHU will likely inflict some degree of psychological trauma upon most inmates confined there for more than brief periods. Clearly, this impact is not to be trivialized; however, for many inmates, it does not appear that the degree of mental injury suffered significantly exceeds the kind of generalized psychological pain that courts have found compatible with Eighth Amendment standards.[33]

Henderson did find that for the group of prisoners who were mentally ill or had a history of prior psychiatric problems, placement in the SHU did constitute an Eighth Amendment violation.

> For these inmates, placing them in the SHU is the mental equivalent of putting an asthmatic in a place with little air to breathe. The risk is high enough, and the consequences serious enough, that we have no hesitancy in finding that the risk is plainly "unreasonable." Such inmates are not required to endure the horrific suffering of a serious mental illness or major exacerbation of an existing mental illness before obtaining relief.[34]

Almost twenty years later, in 2011, thousands of prisoners in California went on a hunger strike protesting the conditions at the Pelican Bay SHU and other SHUs around the state. That hunger strike garnered national and international attention and eventually led to a class action lawsuit claiming that incarceration at Pelican Bay for more than ten years was cruel and unusual punishment in violation of the Eighth Amendment.[35] Some of the same prisoners who were at Pelican in the early 1990s were still there in 2011 and were named plaintiffs in the new class action lawsuit.[36]

California responded to the lawsuit by arguing that Judge Henderson had already ruled that the type of psychological pain and suffering that the ordinary, non–mentally ill prisoner suffered at Pelican Bay did not rise to the level of a constitutional violation, and that only harm that resulted in serious mental illness or attempted suicide would be actionable. None of the ten named plaintiffs in the new *Ashker v. Governor* were mentally ill, although they all claimed serious psychological harm. Moreover, they argued that Judge Henderson's ruling had been based on a record of prisoners who had spent two to three years at Pelican Bay, and that he had specifically left open the possibility that more prolonged stays in solitary confinement might violate the Constitution. Henderson could "not begin to speculate on the impact that Pelican Bay SHU conditions may have on inmates confined in the SHU for periods of 10 or 20 years or more; the inmates studied in connection with this action had generally been confined to the SHU for three years or less."[37] Judge Claudia Wilken, who was assigned to hear *Ashker*, rejected California's motion to dismiss the lawsuit, finding it was not precluded by Judge Henderson's decision in *Madrid v. Gomez*.[38]

While *Ashker* proceeded, the plaintiffs still faced the substantial hurdle set by Henderson and other cases that generalized psychological pain such as depression, paranoia, lack of concentration or memory, anger, and hallucinations was insufficient, at least if suffered for only several years, to constitute cruel and unusual punishment. The plaintiffs' team had included top notch psychological experts, one of whom, psychologist Craig Haney, had also testified in the *Madrid* case. Moreover, the plaintiffs' psychological harms seemed even more profound than those recognized in *Madrid* and, the team felt, ought to have been sufficient to establish an Eighth Amendment violation. Nevertheless, the law's general discounting of psychological harm and the Supreme Court's reluctance to recognize familiar modes of punishment as cruel and unusual precluded complacency.

The law concerning prisoners, like the torts jurisprudence discussed above, tends to discount psychological pain and suffering, as did Judge Henderson. While the courts have recognized that psychological harm inflicted by prison officials can constitute an Eighth Amendment violation, Congress enacted a statute, the Prison Litigation Reform Act, that precludes prisoners who suffer constitutional violations from being awarded damages unless they can show that they have suffered

"physical injury" and not purely mental harm.[39] Thus, for example, the Eleventh Circuit Court of Appeals dismissed a damages claim in which prison officials had "ordered prisoners to strip naked, and performed body cavity searches while members of the opposite sex were present; . . . made harassing comments to an inmate because of his perceived sexual orientation; and ordered one prisoner to 'tap dance' while naked."[40] So too, while some courts have held that rape or other sexual assaults constitute a physical injury within the meaning of the Prison Litigation Reform Act, several courts have held that "the bare allegation of sexual assault" does not constitute a physical injury under the statute.[41] Furthermore, when the Senate ratified the Convention on the Prevention of Torture, it added a reservation that mental harm would not count as torture unless it fell within certain narrowly circumscribed exceptions.[42] As it does with tort law, the United States treats mental pain as a second-class citizen for purposes of the international law of torture.

Given the reluctance of the courts and Congress to fully recognize that the mental pain wrought by solitary confinement rises to the level of an Eighth Amendment violation, plaintiffs' counsel sought ways of bringing other sciences and social sciences to demonstrate the harm caused by such conditions. In this case, the science was brought to bear in support of a conclusion that seemed obvious. To hold a person in a small cell with no windows for twenty-three hours a day under crushing conditions of isolation for ten, fifteen, or twenty years must cause serious harm to that individual in a manner that civilized society should not tolerate. As one prominent court of appeals judge has noted, it seems "pretty obvious, that isolating a human being from other human beings year after year or even month after month can cause substantial psychological damage, even if the isolation is not total."[43] Or as Justice Kennedy wrote in a concurring opinion in a case that did not directly challenge the use of solitary confinement, "the human toll wrought by extended terms of isolation has long been understood and questioned by writers and commentators. . . . [R]esearch still confirms what this Court suggested over a century ago. Years on end of near total isolation exact a terrible price."[44]

The plaintiffs' use of neuroscience in the solitary confinement challenge was thus similar to the role neuroscience played in the Eighth Amendment challenge to the execution of juveniles, wherein the Court viewed scientific evidence not as an independent basis for decision, but as evidence that would tend to confirm the conclusion that prolonged solitary confinement caused serious mental and physical harm to the brain to a degree prohibited by the Constitution. As the Court noted in the juvenile death penalty case *Roper v. Simmons*, in distinguishing between adults and juveniles, "as any parent knows, and as scientific and sociological studies respondent and his amici cite tend to confirm, 'a lack of maturity and underdeveloped sense of responsibility are found in youth more often than in adults, and are more understandable among the young.'"[45]

Using neuroscience in the prisoner context, however, faced substantial obstacles. The most important was that neuroscientists had never studied the brains of prisoners and, therefore, no studies directly on point existed. Moreover, the possibility that neuroscientists could do significant scientific studies of the Pelican Bay prisoners was remote. To demonstrate conclusively that solitary confinement alters the brain, a study would have to use one of two types of design. The optimal design would be longitudinal and would require gathering baseline brain imaging data on prisoners before they were placed in solitary confinement followed by periodic testing

to ascertain changes in brain structure and function. To be certain that such changes were associated with isolation and not with prison life in general, similar observations of well-matched control subjects (of similar age, sex, mental ability, and ideally criminal offense history) would have to be taken over the same period of time. An additional control group of subjects equally well-matched on crucial variables but not incarcerated would also be useful since this would enable the parsing of the effects of the general stress of prison life from the additional impact of social isolation, physical inactivity, and other distresses of solitary housing. Absent the basal data, a less optimal cross-sectional design could be used, but it would require a larger number of prisoners in order to enable either the two-way or three-way comparison.

Not only would the cost of doing such a study be massive and untenable for a public interest lawsuit, but even if the necessary funds could be raised, prison officials do not allow scientists into the prison to do studies, and, absent an unlikely court order, the plan would not be workable. Thus, using neuroscience to aid the Court in understanding how prolonged solitary confinement affected the brain required drawing on extant knowledge and theory and extrapolating from what scientists know generally about the brain to the situation in which these prisoners found themselves. This is a second-best solution, but the lawyers thought it would be nonetheless valuable to the Court, even though a more definitive study of the type sketched above was not possible for the purposes of *Ashker v. Brown*.

Despite these obstacles, the *Ashker* lawyers decided to make neuroscience evidence part of their core case for two reasons. First, the Supreme Court has held that to establish an Eighth Amendment violation, a prisoner must show that he or she has been deprived of some basic human need such as food, sleep, or exercise.[46] Court challenges to solitary confinement have sought to add social interaction to the list of basic human needs, and in some cases, have been successful.[47] Neuroscience could aid in establishing that the human brain requires social interaction with other people and, therefore, such interaction is a basic human need. In *Ashker*, the plaintiffs submitted an expert report from neuroscientist Matthew Lieberman, the director of the Social Cognitive Neuroscience Laboratory at the University of California, Los Angeles, and author of the award-winning book, *Social: Why Our Brains Are Wired to Connect*.[48] His declaration explained why social interaction is a basic human need on a par with sleep or exercise. The deprivation of that human need will not – unlike the deprivation of food – result in death in a short order, but like the deprivation of sleep or exercise, it will have very deleterious effects on both mental and physical health over time.

The second reason to introduce neuroscience evidence was to break down the divide between mental and physical pain. The research suggests that solitary confinement would produce physiological changes in the brain, harm that is therefore physical, potentially observable, and causes mental pain. As in the tort context, a demonstration of physiological harm would supplement the psychological research of the harm suffered by individuals who are denied social contact.

Ashker is but one of several cases in which neuroscience has been used to challenge prolonged solitary confinement. As already mentioned, the *Ashker* plaintiffs introduced Lieberman's expert report to support their claims that solitary confinement causes serious mental and physical harms and deprives those confined of the basic human need of social interaction. Lieberman had never studied prisoners nor solitary confinement in state prisons, but he

applied his general research on the effects of social isolation on the brain to the Pelican Bay context.

Lieberman started his report with the proposition that "it is considered settled science within the field of psychology that humans and all mammals have a fundamental need for social connection."[49] Lieberman then described the neuroscientific contribution to understanding social connection as a basic need. He summarized that

> the brain has a neural system that registers various kinds of physical pain – each linked to a potential survival threat (loss of food, water, shelter). . . . My lab and others have observed that when individuals are in a socially deprived state, they experience *social pain* and this produces neural activity consistent with it being a form of pain.[50]

To Lieberman, his neuroscience research, along with the work of others, provides compelling evidence that the social pain of isolation involves "the same neural and neurochemical processes invoked during physical pain."[51] Indeed, fMRI studies that he conducted in collaboration with psychologist Naomi Eisenberger demonstrated that when people were subjected to social isolation, it affected neural activity in certain cortical regions of the brain associated with physical distress, in the same way physical pain would. Lieberman's study has been replicated dozens of times in labs around the world. Lieberman concluded that the social pain caused by isolation is not metaphorical pain, but has a physical effect on brain activity causing the brain to signal distress.[52]

The Amicus Curiae Brief of Medical and other Scientific and Health Related Professionals filed in the United States Supreme Court case of *Ziglar v. Abbasi* also used neuroscience studies to support the proposition that solitary confinement causes both serious psychological and physical harm.[53] The brief cites coauthor Huda Akil for the proposition that neuroscience studies suggest that solitary confinement can "fundamentally alter the structure of the human brain in profound and permanent ways."[54] Akil's view reflects the knowledge that the human brain, like all mammalian brains, alters its structure and functioning based on stimuli from its environment. This process, termed "neuroplasticity," subsumes several mechanisms, including changes in branching or arborization of neurons to enable new connections to neighboring brain cells, changes in activity of certain brain circuits, and, in specialized brain regions, changes in the rate of birth of new neural cells that become embedded in critical circuits.

One region that is very "plastic" is the hippocampus (or seahorse, due to its shape). The hippocampus plays a critical role in handling the interface of the individual with the external world by mapping the physical environment in three dimensions: it sets the level of emotional reactivity and anxiety, it encodes stressful events and controls the body's response to stressors, and it plays a primary role in encoding memories of recent events and determining whether they are destined for long-term storage elsewhere in the brain. These changes are typically adaptive in that they enable the individual to assess a context (physical and emotional), react to it appropriately, and remember it and anticipate future responses. But under conditions of severe and sustained stress, the hippocampus loses this neuroplasticity: it physically shrinks, the rate of birth of new cells diminishes or ceases, the arbors regress, and the opportunity for contacts with neighboring cells decreases. It is therefore not surprising that this brain region begins to fail in its functioning, with loss of emotional and stress control, loss of stress regulation, sometimes defects in memory, spatial orientation, and other cognitive processes, and in extreme cases, last-

ing changes in mood, including severe depression. Moreover, since the brain is highly interconnected, this is but one node of many changes that propagate across the brain and greatly diminish the individual's affective and cognitive functions, resulting in long-term deficits in each.

As argued by Akil in the context of the *amicus* brief, each of the key features of solitary confinement – lack of meaningful interaction with others and the natural world and lack of physical activity and visual stimulation – "is by itself sufficient to change the brain . . . dramatically depending on whether it lasts briefly or is extended."[55] As noted in the brief, many neurobiological studies "reveal that certain regions of the brain of people who experience extreme psychological stress (like those in solitary confinement) literally diminish in volume because the neural cells become shriveled."[56]

A large body of animal studies strongly supports the notion of altered neuroplasticity as a result of an impoverished environment. In a Canadian case, challenging prolonged solitary confinement in British Columbia, the lawyers sought to introduce an expert report from neurologist and animal behavior scholar Michael Zigmond, who noted that the rats and mice that he studies have 99 percent of the same genes as humans and that the basic neuroanatomy of the mouse parallels that of humans.[57] Zigmond reports that his and other studies demonstrate that when mice and rats are randomly grouped into two different environments, one that is enriched with lots of activities and another that is isolated, the rodents in the isolated environment show "enormous differences," such as a "decrease in the anatomical complexity of the brain (including fewer connections between nerve cells and even fewer nerve cells) and a decrease in the number of blood vessels in the brain."[58] These animals also show differences in learning and memory, as well as susceptibility to a range of diseases that emulate human diseases such as Alzheimer's disease, Parkinson's disease, and strokes.[59]

Zigmond concludes that "some of these effects are undoubtedly related to one or more of the biochemical effects of isolation, which include a decrease in the concentration of 'neurotrophic factors' or growth factors that are responsible for the repair of neurons should they begin to atrophy."[60] A key neurotrophic factor is brain-derived neurotrophic factor (BDNF), which modulates diverse functions including learning, memory, navigation, and mood. Similarly, Zigmond has reported that isolation decreases the synthesis of the neurotransmitter dopamine, which is critical for motor function and reward, and the capacity to reduce inflammation and oxidative stress.[61]

Zigmond's most recent and in-depth study showed that brains of isolated rodents have smaller neurons, with fewer branches in the hippocampus and cerebral cortex regions, which affect learning, memory, and executive brain functions.[62] The one region that does show more activity is the amygdala, which mediates fear and anxiety, symptoms reported by human prisoners confined in solitary.

Mice and rats, of course, are not humans, and therefore these studies do not prove that human brains are affected in the same ways as those of rodents.[63] Nonetheless, there are similarities, and the fact that rodents and other mammals react to isolation in a manner that affects their brain functions is some evidence that the human brain is likely to be similarly affected. Thus, this body of work by neuroscientists is not dispositive. But, paraphrasing Justice Kennedy's observation in *Roper v. Simmons*, this research tends to confirm what common experience and years of psychological studies teach us: that prolonged solitary confinement can cause both serious psychological and physiological harm.

One would think it self-evident from a purely ethical perspective that placing a person in a small cell for twenty-three hours a day with very limited or no social contact for years, and sometimes for decades, should not be permitted in civilized society. However, the law requires evidence that such treatment would cause serious harm, and it is in this domain that neuroscience can play an important role in the legal struggle against prolonged solitary confinement. As discussed above, neuroscience is potentially relevant not just to this but to a wide range of other legal issues because an underappreciated and often overlooked contribution that neuroscience can bring to the law is to break down the division that currently exists between physiological and psychological harm and between physical and mental injury. Neuroscience challenges the law's long-unchallenged assumption that most mental suffering is inescapably subjective. Proceeding from the obvious truth that the brain is a physical organ, neuroscience can show empirically and explain theoretically that the brain both regulates and is profoundly affected by mental harm and suffering.

As the interface between neuroscience and the law evolves, several challenges are likely to emerge. While we have underscored the value of neuroscience in providing scientific support for commonsense notions, there will likely be situations in which the opposite happens. Science teaches us that, on occasion, what seem to be obvious truths are incorrect. An example is the widely held belief that children are intrinsically resilient, that they will not remember early life trauma, that they will simply not encode the stress, or that they will readily forget it. However, neurobiological evidence clearly shows that early-life traumatic events, especially if repeated, can produce a lasting deleterious effect on the individual that will manifest later in life. Societal views, as well as legal thought, will likely need to be modified to incorporate such insights.

Moreover, when neuroscience accords with common sense, it may nonetheless provide novel perspectives that may be impactful on legal decisions and legal thought. For example, neuroscience has validated the importance of so-called critical periods during human development when major epigenetic, cellular, and molecular reprogramming can take place in response to environmental conditions, but it has also shown that such key periods are not confined to early childhood. One key period occurs during adolescence. As additional biological evidence accumulates, it will be important for the law to contemplate the implications of such a major biological upheaval, both in understanding human behavior and in dealing with it from a legal standpoint.

Another major challenge stems from the fact that neurobiological changes are rarely binary. Rather, they are incremental, reflecting processes that may wax and wane, and the threshold at which a change becomes deleterious can be difficult to discern. For example, as described above, stress remodels the brain. Some level of remodeling is adaptive and enables coping with further stress, but chronic or severe stress becomes maladaptive, leading to neural damage. However, the point at which a change is likely to be damaging rather than helpful is unclear and varies as a function of the preexisting vulnerability or resilience of any given individual. Moreover, as tools and techniques in neuroscience evolve, our ability to detect changes will improve.

The existence of these continua is not readily compatible with legal formalisms that may classify matters in more binary ways. An example is the notion of competency. As neuroscientists develop more robust biomarkers of cognitive function, it may be possible to detect loss of competency in some functions (such as recall

of recent events) coexisting with maintenance of competency in other brain functions (such as recall of distant events or moral judgment). This may push legal thought toward a more nuanced definition of competency or facets thereof, informed by scientific knowledge.

Our thinking about the ethical, philosophical, and legal implications that arise from the explosion in neuroscience knowledge is in its infancy. It is clear, however, that ongoing discourse between the disciplines will profit both the science and the law, framing questions in interesting ways for the neuroscientist and challenging legal professionals to amend old or develop new conceptual frameworks.

ENDNOTES

[1] Jeffrey Rosen, "The Brain on the Stand," *The New York Times Magazine*, March 11, 2007; Greg Miller, "The Brain Gets Its Day in Court," *The Atlantic*, March 1, 2016; and Nita A. Farahany, "Neuroscience and Behavioral Genetics in U.S. Criminal Law: An Empirical Analysis," *Journal of Law and the Biosciences* 2 (3) (2015): 485–509.

[2] Owen D. Jones, René Marois, Martha J. Farah, and Henry T. Greely, "Law and Neuroscience," *The Journal of Neuroscience* 33 (45) (2013): 17624–17630.

[3] Ibid. See also Owen D. Jones, Jeffrey D. Schall, and Francis X. Shen, eds., *Law and Neuroscience* (New York: Aspen Publishers, 2014).

[4] Shaun Cassin, "Eggshell Minds and Invisible Injuries: Can Neuroscience Challenge Longstanding Treatment of Tort Injuries?" *Houston Law Review* 50 (3) (2013): 929–962; Betsey G. Jones, "The Future of Emotional Harm," *Fordham Law Review* 83 (5) (2015): 2605–2653 [neuroscientific demonstrations that emotional injury occurs with physiological injuries to the brain have begun to impact the legal landscape]; and *Allen v. Bloomfield Hills School District* 760 N.W. 2d 811, 812 (Mich. Ct. App. 2008) [using neuroscientific evidence (a PET scan) to hold that the tort bodily injury requirement was met by injury to the brain].

[5] See Matthew D. Lieberman, expert report, submitted in *Todd Ashker et al. v. Governor of the State of California et al.*, 4:09-cv-05796-CW (N.D. Cal. 2014), https://ccrjustice.org/sites/default/files/attach/2015/07/Lieberman%20Expert%20Report.pdf. See also Craig Haney, Michael Zigmond, Jules Lobel, and Robert King, "Neuroscience in Policy: Solitary Confinement in California," panel at Law and Neuroscience Conference 2017, University of California, San Francisco, and University of California, Hastings, February 16–17, 2017, http://www.ucconsortium.org/events/law-neuroscience-conference-2017/. *Ashker v. Governor* Case No. 4:09-cv-05796-CW was eventually settled in September 2015 in an agreement to release virtually all of the thousands of prisoners being held in indeterminate solitary confinement in California. Ian Lovett, "California Agrees to Overhaul Its Use of Solitary Confinement," *The New York Times*, September 1, 2015.

[6] Jones et al., "Law and Neuroscience," 17624 [see note 2].

[7] David Faigman, "Review: Law and Neuroscience," *Jurimetrics Journal* 56 (2) (2016): 197–200.

[8] Steven J. Morse, "Criminal Law and Common Sense: An Essay on the Perils and Promise of Neuroscience," *Marquette Law Review* 99 (1) (2015): 39–74.

[9] Erin Murphy, "Neuroscience and the Civil/Criminal *Daubert* Divide," *Fordham Law Review* 85 (2) (2016): 619, 629.

[10] *Roper v. Simmons*, 543 U.S. 551, 569 (2004).

[11] Morse, "Criminal Law and Common Sense" [see note 8].

[12] Murphy, "Neuroscience and the Civil/Criminal *Daubert* Divide," 630 [see note 9].

[13] David L. Faigman, Christopher Slobogin, and John Monahan, "Gatekeeping Science: Using the Structure of Scientific Research to Distinguish Between Admissibility and Weight in Expert Testimony," *Northwestern Law Review* 110 (4) (2016): 859 – 904.

[14] Ibid., 887 – 890.

[15] Betsey J. Grey, "Neuroscience and Emotional Harm in Tort Law: Rethinking the American Approach to Free-Standing Emotional Distress Claims," *Law and Neuroscience* 13 (2011): 203.

[16] See Jones et al., "Law and Neuroscience," 935 [see note 2]; and Grey, "Neuroscience and Emotional Harm in Tort Law" [see note 15].

[17] *Hawes v. Germantown Mutual Insurance Company*, 309 N.W. 2d 356, 360 (Wis. Ct. App. 1981).

[18] American Law Institute, *Restatement of the Law Third – Torts: Liability for Physical and Emotional Harm* (Philadelphia: American Law Institute, 2009), Section 4, Comment B ["the definition of bodily harm is meant to preserve the ordinary distinction between bodily harm and emotional harm"].

[19] Ibid., chap. 8, scope, sec. 45 – 48.

[20] Karen D. Davis, Herta Flor, and Henry T. Greely, et al., "Brain Imaging Tests for Chronic Pain: Medical, Legal and Ethical Issues and Recommendations," *Nature Reviews Neurology* 13 (2017): 624 – 638; and A. C. Pustilnik, "Imaging Brains, Changing Minds: How Pain Neuroimaging Can Inform the Law," *Alabama Law Review* 66 (5) (2015): 1099 – 1158.

[21] Alexander J. Shackman, Tim V. Salomons, and Heleen A. Slagter, et al., "The Integration of Negative Affect, Pain and Cognitive Control in the Cingulate Cortex," *Nature Reviews Neuroscience* 12 (2011): 154 – 167.

[22] *Allen v. Bloomfield Hills School District* 760 N.W. 811, 812 (Mich. Ct. App. 2008).

[23] Ibid., 815 – 817.

[24] Ibid.

[25] Ibid., 816.

[26] Sarah Baumgartel, Corey Guilmette, and Johanna Kalb, et al., "Time-in-Cell: The ASCA-Liman 2014 National Survey of Administrative Segregation in Prison," Public Law Research Paper No. 552 (New Haven, Conn.: Yale Law School, 2015).

[27] Keramet Reiter, *23/7: Pelican Bay Prison and the Rise of Long-Term Solitary Confinement* (New Haven, Conn.: Yale University Press, 2016).

[28] Complaint in *Todd Ashker et al. v. Governor of the State of California et al.*, 4:09-cv-05796-CW (N.D. Cal., 2014), https://ccrjustice.org/sites/default/files/assets/Ruiz-Amended-Complaint-May-31-2012.pdf.

[29] *Madrid v. Gomez*, 889 F. Supp. 1146 (N.D. Cal. 1995).

[30] Ibid., 1229.

[31] Ibid.

[32] Ibid., 1231 – 1235.

[33] Ibid., 1265.

[34] Ibid., 1265 – 1266.

[35] See, for example, Lovett, "California Agrees to Overhaul Its Use of Solitary Confinement" [see note 5].

[36] In *Ashker v. Governor*, Todd Ashker, Danny Troxell, and Ronald Dewberry were named plaintiffs who had been at the Pelican Bay SHU since the *Madrid* case was tried.

[37] *Madrid v. Gomez* quoted in Reiter, *23/7*, 1257 [see note 27].

[38] Decision and Order on Defendants' Motion to Dismiss in *Todd Ashker et al. v. Governor of the State of California et al.*, 4:09-cv-05796-CW (N.D. Cal., 2014), 10, https://ccrjustice.org/sites/default/files/assets/Order%20Denying%20Motion%20to%20Dismiss%204.9.13.pdf.

[39] Prison Litigation Reform Act of 1995, Pub. L. No. 104-134, § 803(d), 110 Stat. 1321 (1996), codified as 42 U.S.C. § 1997e(e).

[40] *Harris v. Garner*, 190 F. 3.d 1279, 1282 (11th Cir. 1999).

[41] *Hancock v. Payne*, No. Civ.A.1.03CV671JMRJMR, 2006 WL 21751, at *3 (S.D. Miss. January 4, 2006); and *Smith v. Shady*, No. 3:CV-05-2663,2006 WL 314514, at *2 (M.D. Pa. February 9, 2006). See Margo Schlanger and Giovanna Shay, "Preserving the Rule of Law in America's Jails and Prisons: The Case for Amending the Prison Litigation Reform Act," *University of Pennsylvania Journal of Constitutional Law* 11 (1) (2008): 139–154.

[42] U.S. Senate Resolution of Advice and Consent to Ratification of the Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, 138 Cong. Rec. S17486-01 01 (daily ed. October 27, 1990).

[43] *Davenport v. DeRobertis*, 844 F.2d 1310, 1313 (7th Cir. 1988).

[44] *Davis v. Ayala*, 135 S.Ct. 2187, 2209-10 (2005).

[45] *Roper v. Simmons*, 543 U.S. 551, 569 (2004). Emphasis added.

[46] *Wilson v. Seiter*, 504 U.S. 294, 304 (1991).

[47] See, for example, *Wilkerson v. Stadler* 639 F. Supp. 2. 654 (M.D. La. 2007).

[48] Lieberman, expert report [see note 5].

[49] Roy F. Baumeister and Mark R. Leary, "The Need to Belong: Desire for Interpersonal Attachment as a Fundamental Human Need," *Psychology Bulletin* 117 (3) (1995), quoted in Lieberman, expert report, 2 [see note 5].

[50] Ibid., 5.

[51] Ibid., 6.

[52] Ibid., 8–9.

[53] 582 U.S. ___(2017). *Ziglar* challenged the conditions under which certain alien arrestees were held in custody following 9/11, including solitary confinement in tiny cells for more than twenty-three hours a day, often in shackles, for as long as eight months. The case, however, was eventually decided on technical legal grounds relating to whether the defendants could be sued for damages under relevant Supreme Court precedent. Thus, the neuroscience arguments advanced by *amici* played no role in the decision, though the arguments further illustrate the relevance of neuroscientific evidence to litigation over conditions of confinement. Supreme Court of the United Sates, "Brief of Medical and Other Scientific and Health-Related Professionals as *Amici Curiae* in Support of Respondents and Affirmance," *Ziglar v. Abbasi*, No. 15-1358, December 23, 2016.

[54] Supreme Court of the United Sates, "Brief of Medical and Other Scientific and Health-Related Professionals as *Amici Curiae* in Support of Respondents and Affirmance," 24–25 [see note 53].

[55] Kate Allen, "Researchers Study Effects of Prolonged Isolation Among Prisoners," *The Toronto Star*, February 14, 2014, quoted in ibid., 25.

[56] Supreme Court of the United Sates, "Brief of Medical and Other Scientific and Health-Related Professionals as *Amici Curiae* in Support of Respondents and Affirmance," 25 [see note 53].

[57] See Expert Report of Michael J. Zigmond, *British Columbia Civil Liberties Association and the John Howard Society of Canada v. Attorney General of Canada*, SCBC Vancouver Registry No. S-150415 (February 22, 2017), 2 [report was rejected by the Canadian Court and not admitted into evidence]. See also Ramin Skibba, "Solitary Confinement Screws Up the Brains of Prisoners," *Newsweek*, April 18, 2017.

[58] Zigmond, "Isolated Housing of Non-Human Animals," 7 – 8 [see note 57].

[59] Ibid., 8.

[60] Ibid.

[61] Ibid.

[62] Skibba, "Solitary Confinement Screws Up the Brains of Prisoners," 52 [see note 57].

[63] Indeed, the Canadian Court rejected the Zigmond declaration on the grounds that animal-based evidence could not be extrapolated to humans.

Universities: The Fallen Angels of Bayh-Dole?

Rebecca S. Eisenberg & Robert Cook-Deegan

Abstract: The Bayh-Dole Act of 1980 established a new default rule that allowed nonprofit organizations and small businesses to own, as a routine matter, patents on inventions resulting from research sponsored by the federal government. Although universities helped get the Bayh-Dole Act through Congress, the primary goal, as reflected in the recitals at the beginning of the new statute, was not to benefit universities but to promote the commercial development and utilization of federally funded inventions. In the years since the passage of the Bayh-Dole Act, universities seem to have lost sight of this distinction. Their behavior as patent seekers, patent enforcers, and patent policy stakeholders often seems to work against the commercialization goals of the Bayh-Dole Act and is difficult to explain or justify on any basis other than the pursuit of revenue.

REBECCA S. EISENBERG is the Robert and Barbara Luciano Professor of Law at the University of Michigan. She has published extensively in leading legal and scientific journals on patent issues in the life sciences and academic patenting.

ROBERT COOK-DEEGAN is a Professor at the School for the Future of Innovation in Society and Consortium for Science, Policy & Outcomes at Arizona State University. He is the author of *The Gene Wars: Science, Politics, and the Human Genome* (1996) and has published in such journals as *Science*, *New England Journal of Medicine*, and *Nature Biotechnology*.

The Bayh-Dole Act of 1980 established a new default rule that allowed nonprofit organizations and small businesses to own, as a routine matter, patents on inventions resulting from research sponsored by the federal government.[1] The new law replaced divergent and changeable rules and practices of different federal funding agencies. It made ownership more predictable and reduced the need for case-by-case negotiations to secure rights.[2]

University patent ownership featured prominently in subsequent commentary on the Bayh-Dole Act, but the initial choice to limit the new rule to nonprofit organizations and small businesses was a matter of political expediency. Although universities helped get the Bayh-Dole Act through Congress, the primary goal, as reflected in the recitals at the beginning of the statute, was not to benefit universities but to promote the commercial development and utilization of federally funded inventions.[3] It was part of a broader initiative to give patent ownership to research contractors, rather than to federal funding agencies, in order to accelerate commercial development. By ap-

doi:10.1162/DAED_a_00521

plying the new default ownership rule only to nonprofit organizations (including universities) and small businesses, advocates sidestepped decades-old objections to giving patent monopolies to powerful business interests when inventions were made at taxpayer expense. Had large businesses been included in Bayh-Dole, an anonymous Senate aide confided to a reporter, "the bill would never have [had] a chance of passing."[4] But the limitation was only temporary. Soon enough, the new policy was extended to large businesses, first quietly in an executive order signed by President Ronald Reagan, and then more durably in an inconspicuous amendment to the statute.[5]

Universities fit awkwardly in the arguments for patent ownership by contractors rather than government agencies. Advocates emphasized that government agencies do not commercialize inventions themselves, and therefore government ownership inevitably required costly licensing transactions to transfer the rights that firms required to protect investments in commercialization. Government stewards might be cautious about giving public property away to private firms, introducing uncertainty and delay. When the contractor was a private firm, commercialization could proceed more quickly by vesting rights in the contractor from the outset. But universities are quite different from private firms. Like the government, universities do not themselves commercialize inventions, but must license their patents for commercialization to proceed. Moreover, in 1980, most universities were relative newcomers to the patent system, having generally avoided patenting for much of the twentieth century, concerned that patenting conflicted with their mission to disseminate knowledge.[6] Universities had no more, and arguably less, expertise in licensing than the government agencies that were criticized as ineffective, and had a similar history of hostility toward patents.

Universities had, however, another argument for patent ownership: only they could provide the close collaboration between faculty and commercial licensees necessary to achieve effective technology transfer for early-stage inventions made in academic laboratories. Patent ownership would give universities and their faculties incentives to secure patent rights and to aid commercial licensees in developing their inventions and bringing them to market. Otherwise, universities would have little reason to divert time and resources away from their academic missions in order to secure patents and to collaborate with licensees. Universities' history of forsaking patents in favor of publication and the dissemination of knowledge made this account plausible. It also made universities seem more trustworthy than business firms: universities would use their patents for public benefit rather than private gain. The perceived halos over universities lit the path to passage of the Bayh-Dole Act.

As a justification for university patents, the logic of technology transfer has limits. Some university inventions surely fit the paradigm of early-stage discoveries requiring further substantial private investment, assisted by university scientists, to launch as commercial products. An important example highlighted in the Bayh-Dole hearings was candidate drugs funded by the National Institutes of Health (NIH) medicinal chemistry program. Private firms had proven unwilling to develop these drugs and to shepherd them through the FDA approval process under the terms of NIH agreements from the 1960s that restricted the firms' ability to secure exclusive rights.[7] Exclusive patent rights were necessary to motivate pharmaceutical firms to invest in expensive clinical trials of promising new drugs. The NIH responded by developing Institutional Patent Agreements (IPAs) that enabled universities to patent drugs resulting from federally funded re-

search and to license their rights to firms. But some IPAs were stalled in administrative review while Bayh-Dole was pending, fueling university interest in codifying patent ownership rules.

Many university inventions, however, do not require substantial postdiscovery investment and the assistance of faculty inventors to achieve commercial application. Some of the most lucrative university patents covered broad enabling technologies that would have been ready for widespread use with or without university patents. These included the Cohen-Boyer patents on basic recombinant DNA techniques (which have generated approximately $255 million for Stanford University and the University of California) and the Axel patents on methods to introduce genes for foreign proteins in eukaryotes (which have generated approximately $800 million for Columbia University).[8] Such technologies face little risk of languishing in academic archives if they are published without patents. Patenting them may provide revenue for universities, but it does not further Bayh-Dole's explicit goal of promoting the development and dissemination of new technologies. According to Niels Reimers, who developed the licensing scheme for the Cohen-Boyer patents, patents on such platform technologies impose a "tax" on subsequent applications, redounding to the benefit of universities, which then use the funds for education and research.[9] Universities prize such patents as a source of unfettered discretionary funds, but they do not promote commercialization; rather, they make commercial development more costly by imposing a need to negotiate and pay for licenses.

Congress recognized that contractor ownership might not be the best way to achieve its goals in all cases. Bayh-Dole provided several mechanisms to depart from this default rule in the terms of funding agreements. In "exceptional circumstances," the agency could determine that withholding title to the invention would better promote the goals of the Act.[10] An agency could also exercise statutory "march-in rights" to license Bayh-Dole patents if it determined that the university or its exclusive licensee was not taking steps to achieve "practical application of the subject invention" or, if necessary, "to alleviate public health or safety needs."[11] Finally, the government retained a paid-up, nonexclusive license to use or to authorize others to use the inventions on behalf of the government.[12] But federal research sponsors have made little use of these provisions to date, perhaps because of burdensome procedural requirements.[13] These requirements were no accident. The architects of Bayh-Dole sought to overcome hostility toward patents in universities and in some funding agencies that they saw as an obstacle to commercial development.

The statute did not limit the new ownership rule to inventions requiring follow-on investment to promote development. And universities have not imposed such limits on themselves.[14] Universities soon came to regard their Bayh-Dole patents as entitlements, using them to generate revenue even when licensing rights were unnecessary for commercialization. The result may actually impede commercialization in some cases, and certainly makes it more expensive.

The drafters of the Bayh-Dole Act may have failed to realize that antipatent attitudes were quickly declining in the academy. Bayh-Dole accelerated a trend that was well under way in the 1970s to reverse formal policies against patenting and to establish university technology transfer offices.[15] As economist Bhaven Sampat observes, Bayh-Dole fostered university patenting "by providing strong Congressional endorsement for the position that active university involvement in patenting and licensing, far from being ignoble, serves the public interest."[16] Perhaps universities could keep their halos while plow-

ing patent revenues back into research and education.

One could argue for university patent ownership as a way to give universities financial rewards for valuable inventions. Notably, this is not among the seven goals recited in the Bayh-Dole preamble.[17] Such an explicit recital might well have drawn political fire. But the argument was not even made.

Sometimes legislation enacted for one purpose turns out to serve another, equally important purpose. Whatever the intent in 1980, enhancing university revenues through patents may now seem like sound policy. But even now, the revenue-for-universities rationale is raised only *sotto voce*, if at all. In case after case, universities justify their patent rights by appealing to the danger of inventions languishing for lack of patent protection, even when university patents are plainly unnecessary for commercialization. The argument persists because promoting commercialization, not revenue, was the foremost justification for university patent ownership in the Bayh-Dole Act.

The overall impact of Bayh-Dole has been a topic of lively debate in the thirty-eight years since its passage.[18] University patenting has dramatically increased, and a few universities have made a lot of money from royalties. Yet licensing revenues remain a small portion of university budgets overall. Respondents to a 2015 survey of the Association of University Technology Managers reported $2.5 billion in licensing revenues (including revenues from trademark, copyright, and unpatented technologies). This is less than 4 percent of the $66.6 billion in university research expenditures, with the wealthiest universities capturing most of the benefits. Although university technology transfer professionals take credit for stimulating commercial development of new technologies, it is not clear how much of that development would have occurred without university patents.

In the post–Bayh-Dole era, universities – the third-largest employer of lobbyists – have had some success in getting Congress to shape patent law to favor their interests.[19] They have secured statutory changes that fortify university patents and make it harder for firms to avoid liability for infringing them.

Meanwhile, universities have had important losses in the courts, especially before the U.S. Court of Appeals for the Federal Circuit (Federal Circuit), an intermediate appellate court with consolidated jurisdiction over patent matters. In a growing body of case law, the courts have refused to adapt patent doctrine to accommodate the circumstances of university research, sometimes with open skepticism toward arguments that universities are acting in the public interest. At times, special pleading from universities seems to have backfired, provoking courts to fortify doctrines that limit the patenting of the kinds of early-stage discoveries that universities often produce.

Universities have, in many cases, pursued patents that they could enforce against product-developing firms for the evident purpose of getting a piece of the action in lucrative technologies that were already being actively developed without the need for university patents. They have sometimes worn their academic halos to court, seeking to adapt patent doctrine to privilege the interests of universities over the competing interests of product-developing firms. This agenda has met with considerable skepticism from the Federal Circuit, which has sometimes explicitly questioned whether university patents are promoting or impeding commercial product development.

An early sign that universities were pursuing patents that were unnecessary for commercial development was the involvement of universities in interference proceedings – administrative proceedings within

the U.S. Patent and Trademark Office to determine priority of invention – in which the university claimed to have made the invention before a commercial inventor. When a university competes with a near-simultaneous commercial inventor seeking patent rights on the same invention, it is difficult to argue that the university is just trying to preserve incentives for commercialization of an invention that would otherwise fail to attract commercial interest. Interferences are only available to establish priority for patent applications filed before March 16, 2013, when U.S. law changed to award priority to the first inventor to file a patent application rather than to the first to make the invention.[20] But before that date, interferences were especially common in biotechnology, a field in which multiple research teams often compete intensely to reach the same goals (such as cloning an obviously important gene). Health policy scholars Jonathan Merz and Michelle Henry found that interferences in biotechnology and organic chemistry were six times more frequent than for patents on average, and that most of the highly competitive "races" the authors cited involved academic research institutions.[21]

Having to litigate a costly interference against a university can only increase the costs and risks facing a product-developing firm. Yet universities persisted in these costly battles, appealing to the Federal Circuit when they lost in the Patent Office. In a priority dispute between academic patent applicants and a pharmaceutical firm over an assay to identify anticancer compounds, for example, the University of Texas Southwestern Medical School pursued repeated appeals to the Federal Circuit, although presumably the firm's commercialization incentives would have been adequately protected by its own already-issued patent on the same invention.[22]

Another strategy for universities more interested in revenues than in promoting product development is to seek broad patent rights on the basis of preliminary academic research that would allow them to sue private firms that later develop products not disclosed in the university patent applications. The Federal Circuit has been consistently hostile to these efforts, invalidating university patents in a series of decisions that fortified the patent law requirement that a patent application must include a "written description" of the invention.[23]

Regents of the University of California v. Eli Lilly involved the first commercial recombinant DNA product that ever reached the market: human insulin.[24] University of California (UC) researchers, having cloned the rat insulin gene, obtained a broad patent covering recombinant microorganisms with DNA sequences encoding human insulin, mammalian insulin, and vertebrate insulin, although the only sequence they disclosed was for rat insulin.[25] Meanwhile, scientists at Genentech successfully cloned the human insulin gene and produced recombinant human insulin. The pharmaceutical firm Eli Lilly manufactured and distributed the final product, which began to replace the previously used insulin product purified from slaughtered pigs as a treatment for diabetes. The University of California sued Eli Lilly and Genentech for patent infringement. On appeal, the Federal Circuit held the UC patent invalid on the basis of what was then a controversial application of the written description requirement. The Federal Circuit held that the written description in the UC patent disclosure only showed possession of the gene for rat insulin, and because the human insulin gene had a slightly different DNA sequence (because the human insulin protein has a somewhat different amino acid sequence), the patent disclosure was insufficient to support the claims to genes for human insulin and for all vertebrate and mammalian insulins. Eli Lilly and Genentech were therefore free to

market recombinant human insulin without liability to the University of California.

Later cases used the invigorated written description requirement to invalidate broad university patent claims to methods of treatment based on discoveries of metabolic pathways likely to be useful in developing new drugs. In *University of Rochester v. G. D. Searle*, the Federal Circuit invalidated claims that would have allowed the University of Rochester to demand royalties from pharmaceutical firms that had developed any selective Cox-2 inhibitors.[26] Cox-2 inhibitors are anti-inflammatory drugs with fewer gastrointestinal side effects than aspirin.[27] The Federal Circuit invalidated the university's patent on a "method for selectively inhibiting PGHS-2 activity in a human host" for lack of an adequate written description. The inventors developed an assay to identify Cox-2 inhibitors, but did not identify or describe any specific inhibitors. The court explicitly rejected the argument that this holding "will have a significant impact on the continuing viability of technology transfer programs at universities and on the equitable allocation of intellectual property rights between universities and the private sector," noting that "none of the…policy objectives of the Bayh-Dole Act encourages or condones less stringent application of the patent laws to universities than to other entities."[28]

The Federal Circuit was even more emphatic in its *en banc* decision in *Ariad Pharmaceuticals v. Eli Lilly*.[29] Researchers at Harvard and the Massachusetts Institute of Technology (MIT) described the NFκB pathway that explained the mechanisms of action of several blockbuster drugs. As in the Rochester case, the university researchers had not actually found an inhibitory compound, but their patents broadly claimed methods of regulating NFκB activity. The Harvard/MIT patents were licensed exclusively to Ariad Pharmaceuticals, which sued Eli Lilly, developer of the NFκB inhibitors Evista and Xigris. The Federal Circuit seemed to view the university patents as anticipatory poaching of the work of the pharmaceutical industry rather than as essential enablers of commercialization:

> Such claims merely recite a description of the problem to be solved while claiming all solutions to it…leaving it to the pharmaceutical industry to complete an unfinished invention. Ariad complains that the doctrine disadvantages universities to the extent that basic research cannot be patented. But the patent law has always been directed to the "useful Arts," U.S. Const. art. I, §8, cl. 8, meaning inventions with a practical use . . . and universities may not have the resources or inclination to work out the practical implications of all such research, i.e., finding and identifying compounds able to affect the mechanism discovered. That is no failure of the law's interpretation, but its intention.…[The law] limits patent protection to those who actually perform the difficult work of "invention" – that is, conceive of the complete and final invention with all its claimed limitations – and disclose the fruits of that effort to the public.[30]

The Federal Circuit rejected arguments that the "written description" doctrine that it had used to invalidate this and other broad university patents on early-stage discoveries removed incentives for private investment in the commercialization of university inventions. Perhaps these arguments seemed particularly unpersuasive in a lawsuit against a firm that had developed and brought to market two commercial products without the benefit of any protection provided by the university patents.

The practical significance of this line of cases has been partially eclipsed by more recent decisions from the U.S. Supreme Court limiting patentable subject matter. These decisions, which preclude patents on natural products, laws of nature, and phenomena of nature, provide an alternative basis for invalidating university pat-

ents arising from fundamental discoveries about biochemical pathways.[31] Universities participated in *amicus* briefs that unsuccessfully argued against the approach ultimately taken by the Court. The Supreme Court holdings extend beyond the problem of reach-through patents from universities – the cases that gave rise to the robust written description requirement from the Federal Circuit – and call into question the validity of many commercial patents in the life sciences.[32]

Yet both the Federal Circuit's written description requirement and the Supreme Court's patentable subject matter doctrine reflect similar concerns: that broad patents on early research discoveries might hinder science and impede rather than promote applications of those discoveries. Although both written description and patentable subject matter doctrines apply to all patents, they present more of an obstacle to patenting early stage research discoveries from university laboratories than to patenting commercial products. Universities argued that these consequences contravened the purposes of the Bayh-Dole Act, but to no avail. These decisions thus represent significant losses for universities and provide a countervailing narrative to the story that university patents are necessary for commercial development.

In addition to seeking patents that are not necessary to promote commercialization, universities have put revenue goals ahead of commercialization by enforcing their patents in litigation against firms that have already developed successful commercial products without the benefit of university patents.[33]

A recent example that reached the U.S. Supreme Court is *Stanford University v. Roche Molecular Systems*.[34] The patents at issue arose from NIH-funded research performed by Mark Holodniy, a Stanford postdoctoral fellow. Stanford researchers were collaborating with scientists at the private firm Cetus to develop an HIV assay using the polymerase chain reaction (PCR). PCR is an important technology developed at Cetus that later won its inventor, Kary Mullis, a Nobel Prize. Stanford sent Holodniy to Cetus to learn PCR and to work on an HIV assay. Holodniy then returned to Stanford and tested the assay in the clinic with other Stanford inventors before Stanford filed patent applications. Meanwhile, Roche acquired Cetus's PCR patent rights and began manufacturing PCR-based HIV detection kits. After the patents were issued to Stanford, Stanford sought royalties from Roche. When they failed to reach agreement, Stanford sued Roche for patent infringement. On appeal, the Federal Circuit ruled that Roche, rather than Stanford, was the true owner of Holodniy's interest in the patents based upon its technical analysis of the legal effects of the terms of two different agreements: a "Visitor Confidentiality Agreement" that Holodniy signed at Cetus and a "Copyright and Patent Agreement" that he had previously signed at Stanford.[35] The Supreme Court granted review to consider whether the Bayh-Dole Act required a different result.[36]

Stanford made a compelling argument that allocation of ownership to Cetus/Roche contravened the design of the Bayh-Dole Act to give universities (and other contractors) the first option to claim rights in inventions made in federally sponsored research. As Justice Breyer explained in a dissenting opinion, contractor ownership is necessary to ensure compliance with a set of conditions that the Bayh-Dole Act requires be included in research funding agreements to protect the public interest.[37] These include provisions for retention of government licenses, reporting obligations, and restrictions on permissible assignments. These safeguards are lost when inventions are assigned – even inadvertently, as apparently happened in this

case – to third parties not bound by those agreements. Moreover, the Bayh-Dole Act contemplates a clear hierarchy of claims to patent ownership with the contractor first in line, followed by the sponsor, and with inventors allowed to claim ownership only when neither the contractor nor the sponsor objects.[38] Justice Roberts's majority opinion inverts this order by holding that the Bayh-Dole Act applies only to inventions owned by the contractor, and not to inventions that employees fail to assign properly to contractors. Although the majority opinion purports to apply strict textual analysis to the language of the Bayh-Dole Act, it ignores many textual cues about the design of the statute that support a different interpretation.

On the other hand, it is hard to argue that Stanford's assertion of patent rights promoted commercialization. Roche developed the technology commercially years before Stanford's patents issued. Roche clearly did not rely on Stanford's patents. The patents did not help Roche, but rather gave Stanford an opportunity to claim a share of the proceeds.

Hard cases make bad law. To the extent that *Stanford v. Roche* calls into question whether universities hold secure title to the patents they are trying to license, it jeopardizes the commercialization goals of the Bayh-Dole Act as well as the protections for the public interest that the Bayh-Dole Act addresses in the terms of funding agreements. Stanford's overreaching in this particular case, where university patents were unnecessary for commercialization, may threaten future rights where clear university title is essential for further development.

In 2008, intellectual property law scholar Mark Lemley posed the provocative question, "are universities patent trolls?"[39] The idea that universities can be patent trolls (that is, patent assertion entities that do not themselves commercialize technology but profit by asserting patents against commercial firms) soon became more commonplace, as major research universities used their patents to collect hundreds of millions of dollars in damage awards and settlements.[40] Criticized for behaving like patent trolls, universities have sometimes sought to avoid the reputational costs of litigation by selling their rights to undisputed patent trolls.[41]

Like patent assertion entities, universities can enforce their patents with little fear of provoking counterclaims for infringement of the patents held by the defendants. Patent infringement litigation against universities and academic researchers is quite rare. This allows university scientists to infringe patents in their laboratories with relative impunity even as universities enforce their patents against other institutions.[42] But this is largely the result of forbearance by patent owners rather than legal immunity from suit.

Universities lost a claim to special status as infringement defendants in the case of *Madey v. Duke University*.[43] Physicist John Madey sued Duke for using his patented field electron laser in a university laboratory. Rejecting Duke's argument that the noncommercial character of academic work precludes infringement liability, the Federal Circuit held that the university would be liable for any use that was in keeping with the "legitimate business" of the university:

> For example, major research universities, such as Duke, often sanction and fund research projects with arguably no commercial application whatsoever. However, these projects unmistakably further the institution's legitimate business objectives, including educating and enlightening students and faculty participating in these projects. These projects also serve, for example, to increase the status of the institution and lure lucrative research grants, students, and faculty.[44]

In a footnote, the court added that "Duke …like other major research institutions of higher learning, is not shy in pursuing an aggressive patent licensing program from which it derives a not insubstantial revenue stream."[45] In other words, to the Federal Circuit, universities are not angels entitled to a privileged status in the patent system, but rather a particular kind of worldly institution pursuing its own objectives, including money.

Other academics may have more to gain than commercial firms from suing academic institutions for patent infringement. Indeed, *Madey v. Duke* was a lawsuit brought by a faculty member against his former university. In another currently pending case, the University of South Florida has sued both the NIH and the nonprofit Jackson Laboratories for making and distributing transgenic mice that are used in Alzheimer's disease research.[46] The cases may or may not succeed, but the fact that a university would bring these lawsuits suggests a decline of academic sharing norms as universities seek to profit from their patents.

Universities have also sought to expand their patent rights by lobbying Congress to change the patent laws in their favor, with mixed results.

This strategy backfired in a campaign by Columbia University to extend the term of its lucrative Axel patents. Columbia worked through Senator Judd Gregg of New Hampshire, a Columbia alumnus, who introduced three different bills in an attempt to extend Columbia's patent term.[47] The patents were then under license to multiple commercial firms, none of which stood to benefit by prolonging their royalty obligations to Columbia. When Senator Gregg's backroom legislative maneuvers became public, there was a strong backlash against both him and Columbia from drug manufacturers, consumer groups, and other members of Congress. Senator Gregg responded that Columbia was "a poor little university" contending with "a fair amount of greed on the part of the drug companies."[48] This phrase came back to haunt Columbia when some of its licensees sued to invalidate one of Columbia's patents. During a hearing in that case, the District Court judge, observing eight lawyers for Columbia in his courtroom, quipped "I thought Columbia was a nonprofit organization who couldn't afford this litigation."[49] In 2004, Columbia signed covenants not to sue the companies for infringement of the disputed patent, and later that year further agreed not to sue anyone else and backed away from demanding royalties.[50]

Other university lobbying efforts have been more successful, leading to statutory changes that make it easier for universities to obtain and enforce patents.[51] Some of these moves have been broadly congruent with the goals of the Bayh-Dole Act. The CREATE Act of 2004, for example, facilitates university-industry research collaborations by extending the benefit of a statutory safe harbor that, as originally enacted, prevented the use of nonpublic information as patent-defeating prior art against patent applications filed by other employees within the same firm.[52] As amended by the CREATE Act, the safe harbor also applies to information belonging to another party to a joint research agreement.[53] This is consistent with the objectives of the Bayh-Dole Act. The recitals in the Bayh-Dole Act reflect a clear intent "to promote collaboration between commercial concerns and nonprofit organizations, including universities," and the statutory change facilitates such interactions by allowing free communication in the course of such collaborations without fear of losing patent rights.[54]

Universities had a significant impact on the new first-to-file rules in the America Invents Act of 2011 (AIA). That legislation changed U.S. law to conform to patent laws of other countries by shifting from the in-

vention date to the application filing date as the time for determining whether an invention is patentable in light of the prior art. Universities initially opposed this change. They worried that the first-to-file priority rule might force them to incur substantial patent filing costs to preserve priority, and that scientists would be unwilling to defer publication until after patent filing.[55] In the end, universities persuaded Congress to modify the new rule to retain a modified version of a one-year "grace period" from prior U.S. law.[56] The grace period gives inventors a year after public disclosure before they lose the right to file patent applications.

Although the grace period is formally available to all inventors, it is most likely to benefit universities. Commercial firms that plan to seek patent rights in other countries are unlikely to rely on it, because public disclosures in the United States would defeat their patent rights elsewhere. But to the extent that it facilitates early publication of research results, the grace period may encourage prompt dissemination of new knowledge.

It is harder to identify a public policy argument, however, for changes that universities secured to a "prior user" infringement defense in the AIA. Like the modified grace period, this provision grew out of university resistance to a proposed change in the law. In 1999, following unexpected decisions of the Federal Circuit upholding patents on methods of doing business, Congress enacted a new "prior user" infringement defense, initially available only against business method patents.[57] This defense protected a user who, acting in good faith, completed the invention at least one year before the patent filing date and commercially used it before the filing date.

The AIA expanded the prior user defense in several ways. It broadened it to cover all patents, not just patents on business methods.[58] It extended the defense to certain related parties and assignees of the original prior user. And although it retained the language about "commercial use," in a bow to universities, it added a new provision defining commercial use to include "use by a nonprofit laboratory or other nonprofit entity such as a university or hospital, for which the public is the intended beneficiary." So far so good: expanding the prior user defense to include universities was entirely consistent with the goals of Bayh-Dole.

More troubling, however, was a change that effectively eliminated prior user rights as a defense to infringement of university patents. In response to university lobbying, Congress added the so-called university exception. Under that exception, a defendant may *not* invoke the prior user defense if the invention "was, at the time the invention was made, owned or subject to an obligation of assignment to . . . an institution of higher education."[59] In other words, when universities are sued as infringement defendants, they can invoke prior user rights to avoid liability, but when universities assert their patents against others, prior user rights are unavailable to defendants. This turns the commercialization justification for the Bayh-Dole Act upside down. Rather than using their patents to help commercial firms develop early stage academic inventions into useful products, universities (and only universities) may now use their patents to sue firms that are so far ahead of academic scientists that they had already put the invention to commercial use a full year before the university filed a patent application. Moreover, since the "university exception" turns not on current ownership, but on whether there was an obligation to assign at the time the invention was made, the defense remains unavailable even if the university later sells the patent to a patent assertion entity (that is, a patent troll).

The statutory text shows vestigial remnants of a university halo. The expansion

of the prior user defense to university laboratories includes the qualification that the use be one "for which the public is the intended beneficiary."[60] But universities tarnished their halos by persuading Congress to eliminate prior user rights as a defense against university patents.

These successes in getting Congress to give universities special treatment under patent law show that the university technology transfer community has become a force to be reckoned with in patent policy. But they also show universities using their lobbying muscle in unabashed pursuit of their own financial interests rather than broader public interests in the dissemination and utilization of new knowledge.

The Bayh-Dole Act chose universities and small businesses as the first beneficiaries of a broader policy shift that aimed to facilitate the commercialization of inventions made in the course of government-sponsored research. It allocated ownership of patent rights to contractors rather than to government funding agencies. Universities, as traditional champions of free dissemination of new knowledge, were regarded as trustworthy stewards of patent rights for the public benefit, in contrast to the big business contractors who later benefited from the same policy. The focus on universities and small businesses made it easier to pass the legislation. Nonetheless, the clear goal of the Bayh-Dole Act was not to generate revenues for universities on government-sponsored research, but rather to facilitate commercial development of new technologies that needed patent incentives to induce postdiscovery private investment.

In the years since passage of the Bayh-Dole Act, universities seem to have lost sight of this distinction. Their behavior as patent seekers, patent enforcers, and patent policy stakeholders often seems to work against the commercialization goals of the Bayh-Dole Act and is difficult to explain or justify on any basis other than the pursuit of revenue.

Universities do good work, and more revenue allows them to do more of it. But revenues from university-owned patents remain a small source of revenue for universities overall, and in total account for less than 5 percent of universities' research expenditures. The policy question is: when do the benefits of university patents justify the costs? Meanwhile, technology transfer offices, as opposed to faculty, have come to dominate the voice of universities in debates about patent policy. The result is a tail-wagging-the-dog distortion, in which the interests of universities as patent owners may be overwhelming their broader interests in widespread dissemination and utilization of new knowledge for the public benefit.

ENDNOTES

[1] Act of December 12, 1980, Pub. L. No. 960517, 94 Stat. 3015-28, codified as amended in pertinent part at 35 U.S.C. §§ 200 – 211 [commonly known as the Bayh-Dole Act].

[2] Rebecca S. Eisenberg, "Public Research and Private Development: Patents and Technology Transfer in Government-Sponsored Research," *Virginia Law Review* 82 (1996): 1663 – 1727.

[3] 35 U.S.C. § 200.

[4] William J. Broad, "Patent Bill Returns Bright Idea to Inventor," *Science* 205 (1979): 473 – 476.

[5] Ronald Reagan, "Memorandum on Government Patent Policy," Memorandum to the Heads of Executive Departments and Agencies, Pub. Papers 248, February 18, 1983; and Trademark

Clarification Act of 1984 § 501(13), Pub. L. No. 989-620, 98 Stat. 3367, codified as amended at 35 U.S.C. § 210(c).

[6] Bhaven N. Sampat, "Patenting and U.S. Academic Research in the 20th Century: The World before and after Bayh-Dole," *Research Policy* 35 (6) (2006): 772 – 789.

[7] Harbridge House, Government Patent Policy Study, *Final Report for the FCST Committee on Government Patent Policy* (Boston: Harbridge House, Inc., 1968), 2 – 40, discussed in Eisenberg, "Public Research and Private Development," 1682 – 1684 [see note 2].

[8] National Research Council, *Intellectual Property Rights and Research Tools in Molecular Biology* (Washington, D.C.: The National Academies Press, 1997), 40 – 42; and Alessandra Colaianni and Robert Cook-Deegan, "Columbia University's Axel Patents: Technology Transfer and Implications for the Bayh-Dole Act," *Milbank Quarterly* 87 (3) (2009): 683 – 715.

[9] Niels Reimers, "Tiger by the Tail," *Chemtech* 17 (1987): 464 – 471.

[10] 35 U.S.C. § 202(a).

[11] 35 U.S.C. § 203(1)(a) – (c).

[12] 35 U.S.C. § 202(c)(4).

[13] Arti K. Rai and Rebecca S. Eisenberg, "Bayh-Dole Reform and the Progress of Biomedicine," *Law and Contemporary Problems* 66 (2003): 289 – 314; and Barbara M. McGarey and Annette C. Leavey, "Patents, Products & Public Health: An Analysis of the Cell Pro March-In Petition," *Berkeley Technology Law Journal* 14 (1999): 1095 – 1116.

[14] Ian Ayres and Lisa Larrimore Ouellette, "A Market Test for Bayh-Dole Patents," *Cornell Law Review* 102 (2017): 271 – 331.

[15] David C. Mowery, Richard R. Nelson, Bhanven N. Sampat, and Arvids Ziedonis, "The Growth of Patenting and Licensing by U.S. Universities: An Assessment of the Effects of the Bayh-Dole Act of 1980," *Research Policy* 30 (2001): 99 – 119; Elizabeth Popp Berman, "Why Did Universities Start Patenting?" *Social Studies of Science* 38 (6) (2008): 835 – 871; and Elizabeth Popp Berman, *Creating the Market University: How Academic Science Became an Economic Engine* (Princeton, N.J.: Princeton University Press, 2012).

[16] Sampat, "Patenting and U.S. Academic Research in the 20th Century," 776 [see note 6].

[17] "It is the policy and objective of the Congress to use the patent system to promote the utilization of inventions arising from federally supported research or development; to encourage maximum participation of small business firms in federally supported research and development efforts; to promote collaboration between commercial concerns and nonprofit organizations, including universities; to ensure that inventions made by nonprofit organizations and small business firms are used in a manner to promote free competition and enterprise; to promote the commercialization and public availability of inventions made in the United States by United States industry and labor; to ensure that the Government obtains sufficient rights in federally supported inventions to meet the needs of the Government and protect the public against nonuse or unreasonable use of inventions; and to minimize the costs of administering policies in this area." Act of December 12, 1980, § 200 [see note 1].

[18] For a thoughtful review and synthesis of the literature, see Rochelle Cooper Dreyfuss, "Double or Nothing: Technology Transfer under the Bay-Dole Act," in *Business Innovation and the Law: Perspectives from Intellectual Property, Labour, Competition and Corporate Law*, ed. Marilyn Pittard, Ann L. Monotti, and John Duns (Cheltenham, United Kingdom: Edward Elgar, 2013), 52 – 73.

[19] Brody Mullins, Douglas Belkin, and Andrea Fuller, "Colleges Flex Lobbying Muscle," *The Wall Street Journal*, November 8, 2015.

[20] Leahy-Smith America Invents Act, Pub. L. No. 112-29, 125 Stat. 284, codified as amended in pertinent part at 35 U.S.C. §§ 102, 103. See U.S. Patent and Trademark Office, *Manual of Patent Examination Procedures* (MPEP) § 2159.

[21] John F. Merz and Michelle R. Henry, "The Prevalence of Patent Interferences in Gene Technology," *Nature Biotechnology* 22 (2004): 153–154.

[22] *Brown v. Barbacid*, 276 F.3d 1327 (Fed. Cir. 2002); *Brown v. Barbacid*, 436 F.3d 1376 (Fed. Cir. 2006); and U.S. Patent No. 5,185,248.

[23] 35 U.S.C. § 112.

[24] 119 F.3d 1559 (Fed. Cir. 1997).

[25] U.S. Patent No. 4,652,525. Rat insulin has one more amino acid than human insulin and is not biologically active in humans; indeed, it provokes an immune response.

[26] 358 F.3d 916 (Fed. Cir. 2004).

[27] U.S. Patent No. 6,048,850.

[28] 358 F.3d at 929.

[29] 598 F.3d 1336 (Fed. Cir. 2010).

[30] *Ariad v. Eli Lilly*, 598 F.3d at 1353. See also Pauline Newman, additional views, noting that the subject matter of the patent "is indeed basic research, which was taken to the patent system before its practical application was demonstrated." 598 F.3d, 1358.

[31] *Mayo Collaboration Services v. Prometheus Laboratories*, 566 U.S. 66 (2012); and *Association for Molecular Pathology v. Myriad Genetics*, 133 S. Ct. 2107 (2013). See also *Bilski v. Kappos*, 561 U.S. 593 (2010); and *Alice v. CLS Bank*, 134 S. Ct. 2347 (2014) [reinvigorating traditional exclusion from patentable subject matter for abstract ideas].

[32] Rebecca S. Eisenberg, "Diagnostics Need Not Apply," *Journal of Science and Technology Law* 21 (2) (2015): 256–286.

[33] Jacob H. Rooksby, "University Initiation of Patent Infringement Litigation," *The John Marshall Review of Intellectual Property Law* 10 (623) (2011).

[34] 563 U.S. 776 (2011).

[35] 583 F.3d 832, 841–842 and 844 n. 1 (2009).

[36] 563 U.S. at 784 n. 2.

[37] 563 U.S. at 794.

[38] 35 U.S.C. § 202(d). "If a contractor does not elect to retain title to a subject invention in cases subject to this section, the Federal agency may consider and after consultation with the contractor grant requests for retention of rights by the inventor subject to the provisions of this Act and regulations promulgated hereunder."

[39] Mark A. Lemley, "Are Universities Patent Trolls?" *Fordham Intellectual Property, Media, and Entertainment Law Journal* 18 (611) (2008).

[40] Daniel Engber, "In Pursuit of Knowledge, and Profit: How Universities Aid and Abet Patent Trolls," *Slate*, May 7, 2014, http://www.slate.com/articles/technology/history_of_innovation/2014/05/patent_trolls_universities_sometimes_look_a_lot_like_trolls.html; and Andrew Chung, "Schools that Sue: Why More Universities File Patent Lawsuits," Reuters, September 15, 2015, http://www.reuters.com/article/university-patents-idUSL1N11G2C820150915.

[41] Chung, "Schools that Sue" [see note 40]; Walter D. Valdivia, "Patent Infringement Suits Have a Reputational Cost for Universities," TechTank blog, November 10, 2015, https://www.brookings.edu/blog/techtank/2015/11/10/patent-infringement-suits-have-a-reputational-cost-for-universities-2/; Andrew K. Cordova and Robin Feldman, "Universities and Patent Demands," *Journal of Law and Biosciences* 2 (3) (2015): 717–721; and Paul Basken, "Under Financial Pressure, Universities Give Patent Buyers a Closer Look," *Chronicle of Higher Education*, October 25, 2013, http://www.chronicle.com/article/Under-Financial-Pressure/142613/.

[42] For a review of empirical studies, see Rebecca S. Eisenberg, "Noncompliance, Nonenforcement, Nonproblem? Rethinking the Anticommons in Biomedical Research," *Houston Law Review* 45 (2008): 1059 – 1099.

[43] 307 F.3d 1351, cert. denied, 539 U.S. 958 (2003).

[44] Ibid., 1362.

[45] Ibid., 1362, n. 7.

[46] Frances McMorris, "USF Files Patent Suit in Dispute over Genetically Modified Mice," *Tampa Bay Business Journal*, December 23, 2015, http://www.bizjournals.com/tampabay/news/2015/12/23/usf-files-patent-suit-in-dispute-over-genetically.html.

[47] Colaianni and Cook-Deegan, "Columbia University's Axel Patents," 701 – 706 [see note 8].

[48] Paul Kane, "Hatch, Gregg Sparring over Patent for Columbia U," *Roll Call*, May 22, 2000.

[49] Bernard Wysocki Jr., "College Try: Columbia's Pursuit of Patent Riches Angers Companies," *The Wall Street Journal*, December 21, 2004.

[50] Colaianni and Cook-Deegan, "Columbia University's Axel Patents," 705 [see note 8].

[51] Peter Lee, "Patents and the University," *Duke Law Journal* 63 (2013): 1 – 87, 63 – 75.

[52] Cooperative Research and Technology Enhancement Act of 2004, Pub. L. No. 108-453, 118 Stat. 3596.

[53] This rule has been superseded by the shift to a first to file system under the America Invents Act, but prior to that time it was codified at 35 U.S.C. § 103(c).

[54] 35 U.S.C. § 200.

[55] Lee, "Patents and the University," 68 [see note 51].

[56] 35 U.S.C. § 102(b),(d) (pre-AIA version). Under these provisions, an inventor would lose the right to a patent by waiting more than a year to file a patent application after the occurrence of certain events, including publication, public use or sale of the invention, and patenting the invention in another country.

[57] *State Street Bank & Trust v. Signature Financial Group*, 149 F.3d 1368 (Fed. Cir. 1998); and American Inventors Protection Act of 1999, Pub. L. No. 106-113, tit. IV, 113 Stat. 1501, codifed as amended in pertinent part at 35 U.S.C. § 273.

[58] Codified at 35 U.S.C. §§ 273(a), 372(e)(1)(A),(B), and 273(c)(2).

[59] 35 U.S.C. §§ 273(e)(5)(A).

[60] 35 U.S.C. §§ 273(c)(2).

The Intractability of Inaccurate Eyewitness Identification

Jed S. Rakoff & Elizabeth F. Loftus

Abstract: Inaccurate eyewitness testimony is a leading cause of wrongful convictions. As early as 1967, the U.S. Supreme Court recognized this danger, but the tests it promulgated to distinguish reliable from unreliable eyewitness testimony were based largely on surmise. More recently, substantial research has demonstrated that, while significant improvements can be made in the manner in which lineups, photo arrays, and other identification procedures are conducted, inherent limitations of human perception, memory, and psychology raise, in many cases, intractable barriers to accurate eyewitness testimony. Where barriers to accurate eyewitness testimony exist, one response is to sensitize jurors to the limitations of eyewitness identifications, but studies to date have not shown that special jury instructions can accomplish that purpose. Moreover, research on expert testimony has produced mixed results, with some studies showing that it helps jurors discriminate between good and bad eyewitness evidence, and other studies showing that it merely creates overall skepticism.

JED S. RAKOFF, a Fellow of the American Academy since 2013, is a United States District Judge for the Southern District of New York and a Professor at Columbia Law School.

ELIZABETH F. LOFTUS, a Fellow of the American Academy since 2003, is Distinguished Professor at the University of California, Irvine. She holds faculty positions in Psychological Science; Criminology, Law & Society; and the School of Law. She is the author of *Eyewitness Testimony: Civil & Criminal* (5th ed., with James M. Doyle and Jennifer E. Dysert, 2013), *Psychology* (5th ed., with Camille B. Wortman and Charles A. Weaver, 1999), and *The Myth of Repressed Memory: False Memories and Allegations of Sexual Abuse* (with Katherine Ketcham, 1994).

Most people have never heard of Kirk Bloodsworth. We have. In 1984, Bloodsworth was convicted and sentenced to death for the rape and murder of a nine-year-old girl in Baltimore. No physical or circumstantial evidence linked Bloodsworth to the crime, but no fewer than five eyewitnesses placed him with the victim and/or at the scene of the crime at about the time that the rape and murder were thought to have occurred. Bloodsworth was, in fact, innocent, as DNA evidence later established. The five eyewitnesses had each "fingered" the wrong guy. After nine years on death row, Bloodsworth was set free. Several years later, the actual murderer confessed, and Bloodsworth was formally exonerated. Bloodsworth's plight is more common than many might think.

Since 1989, more than two thousand wrongly convicted persons have been exonerated in state and federal courts. Commonly contributing to and sometimes clearly causing these wrongful convictions are inaccurate eyewitness identifications. Thus, the In-

doi:10.1162/DAED_a_00522

nocence Project, through DNA testing, has now achieved the legal exoneration of more than 340 persons wrongly convicted of very serious crimes, mostly murder and rape. In roughly three-quarters of those cases, inaccurate eyewitness identifications were a material part of the evidence leading to the convictions. In the words of the Innocence Project, eyewitness identification "is the greatest contributing factor to wrongful convictions proven by DNA testing."[1]

This leads to three questions: Why does eyewitness identification evidence play such an important role in our criminal justice system? Why is such evidence so often inaccurate? What can be done about it?

In the United States, the police are called upon to investigate several million crimes each year. Despite improvements in police techniques, an estimated eighty thousand of these crimes are "solved" each year by strangers who witnessed the crimes, known as "eyewitnesses." (We exclude from the term "eyewitness" those who witness crimes committed by people they already know, such as friends or family members.) These identifications sometimes lead implicated suspects to confess and plead guilty. When this does not happen and guilt is contested, these stranger identifications often are key to convincing the police and prosecutors that they have caught the culprit, and they are crucial in persuading judges and juries to convict. Even in the many criminal cases that never make it to trial, the existence of an eyewitness identification of a defendant typically increases both the severity of what the prosecutor will offer by way of a plea bargain and the pressure the defendant's own attorney will bring to bear in urging his client to accept a plea bargain rather than risk a trial.

Why is this so? To begin with, eyewitness testimony is generally simple, straightforward, and powerful. It usually goes something like this:

> Prosecutor: "Do you see anywhere in this courtroom the man you saw assault your neighbor?"
>
> Eyewitness: "Yes – it is that man [pointing to the defendant]."
>
> Prosecutor: "How confident are you that that is the man who assaulted your neighbor?"
>
> Eyewitness: "Absolutely confident – I'll never forget that face."

Unlike accomplice witnesses, the typical eyewitness to a crime is a passerby who has no motive to lie. Unlike circumstantial evidence, eyewitness testimony is directly probative of guilt and frequently expressed with a high degree of certainty. Unlike expert testimony, eyewitness testimony is immediately understood by even the most confused, inattentive, or ignorant juror. And unlike many other kinds of evidence, eyewitness testimony is rarely the subject of any special cautionary instructions from the judge (though, as discussed below, this is beginning to change).

Put differently, the typical eyewitness is someone with whom the typical juror – or for that matter, the typical police person, prosecutor, and judicial officer – can easily identify: an unfortunate passerby who happened to witness a horrific incident that riveted the passerby's attention and that the passerby, perhaps not without some trepidation, comes forward to report like any good citizen. Who can doubt that she is telling the truth?

Indeed, while there are occasional eyewitnesses (such as accomplices) who have motives to lie, the truthfulness of the typical eyewitness is rarely seriously in doubt. So why are eyewitnesses so often wrong? Until recently, this was largely a matter of speculation. Thus, while the fact of erroneous eyewitness identifications was sufficiently evident that it became the "driving force" behind a series of Supreme Court decisions beginning in 1967, the most the

Court could offer a decade later in summarizing the causes of such errors was that "[t]he witness' recollection of the stranger can be distorted easily by the circumstances [of the viewing] or by later actions of the police."[2]

Nevertheless, the Supreme Court laid out a five-factor test for courts to use in assessing the reliability of eyewitness testimony: 1) the opportunity of the eyewitness to view the suspect at the time of the crime; 2) the witness's degree of attention at the time of the viewing; 3) the amount of time between the witness's viewing of the crime and her first identification of the suspect at the time of a lineup or other identification procedure; 4) the witness's "level of certainty" in the accuracy of her identification at the time of the lineup; and 5) the consistency of the witness's pretrial identifications. These are still the federal standards, though, as we shall shortly see, several of these factors have now been called into question.

This is because we are now the beneficiaries of several decades of serious study on eyewitness identifications and why they are often inaccurate. The results of these efforts present a substantially different picture from the "common sense" assumptions underlying the Supreme Court's earlier tests. To begin with, research has isolated those factors that relate to the eyewitness's own perception, memory, and psychology (so-called estimator variables) and those that relate to the impact on the witness of police and prosecutorial actions (so-called system variables). The distinction is important, because one can change and improve police procedures, but there is little one can do about improving an ordinary human being's ability to accurately perceive and remember.

Turning first to police procedures, sometimes the police investigating a crime will cruise around a neighborhood with an eyewitness, usually shortly after the crime, asking the eyewitness if she sees anyone who resembles the perpetrator. Other times, the police will show the eyewitness a single photo (often a mugshot) of a suspect and ask if that resembles the person she saw commit the crime. Such techniques – loosely grouped together under such rubrics as "show-ups" – have all sorts of problems, but since they are mainly used to advance an investigation, rather than to form the basis of an in-court identification, we will largely put them aside for purposes of this essay. Note, however, that a show-up eyewitness who is later asked to be an eyewitness at trial is subject not only to the problems discussed below, but also to such additional problems as "confirmation bias," by which the very fact of the show-up identification predisposes the eyewitness to making the same identification at a lineup or thereafter.

When the police are seeking not just to advance their investigation, but also to obtain identification evidence that can be used in court, they typically make use of a lineup or photo array. In a lineup, a number of individuals (often six or seven) stand side by side, and the eyewitness, who views them from behind a one-way screen, is asked whether any of them is the person whom the witness saw commit the crime. In a photo array (which is much more easily arranged than a lineup and hence is increasingly the technique of choice), the eyewitness is shown a number of photos and again asked (in various formulations) whether any of them is the person whom the witness saw commit the crime.

Both of these techniques have been the subject of considerable study, much of which has centered on how the form of the procedure, or the way in which it is administered, may suggest to the eyewitness that she should select a particular person. Most obviously, it was established early on that a lineup or photo array in which one of the

individuals or photos stood out from the rest often led to misidentifications. Similarly obvious was the biasing that occurred when the police person administering the test said things like "take a hard look at the third photo." But less overt cues can also influence eyewitness choice. Studies strongly suggest that because many eyewitnesses deeply desire to give the "right" answer, even very subtle feedback or other cues from the police person administering the test, such as nods of approval or body language, can substantially influence whether the eyewitness makes an identification and whom the eyewitness selects.

Although less well developed, there is also some indication that prosecutorial suggestiveness occurring subsequent to a lineup or photo array identification – such as when a witness is in the prosecutor's office being prepared for testifying – may increase the witness's level of confidence in her identification. Somewhat ironically, while defense counsel are often present when the police conduct a lineup or photo array, or, even if not present, can obtain at least some record of what occurred, eyewitness preparation in a prosecutor's office is a largely secret, *ex parte* affair, about which a defense lawyer can only inquire speculatively.

We suggest below some ways in which these system variables can be controlled in order to minimize suggestiveness. But far less tractable are the "estimator variables." To begin with the obvious, an eyewitness's ability to perceive accurately the people and circumstances surrounding the commission of a crime is materially affected by lighting, by distance and angle, by the eyewitness's eyesight, and by the amount of time the eyewitness had the opportunity to view the perpetrator, among other factors. Judges and jurors, as well as police and prosecutors, are generally familiar, through their own experience, with such conditions, and have at least some ability to weigh them. However, many studies indicate that most people regard their own ability to perceive things accurately as much better than it really is, and this may lead them to place greater confidence in an eyewitness's similar ability than is warranted.

Somewhat less obvious is the fact that, as several studies have shown, an eyewitness who encounters a criminal carrying a weapon will often focus more on the weapon than on the face of the perpetrator. Even less obvious are studies (not wholly consistent with each other) suggesting that an eyewitness feeling a modest level of stress at the time of the encounter will perceive it with greater focus, while an eyewitness feeling extreme stress may experience more difficulty in remembering the incident, particularly the peripheral details.

All of this, however, is just the tip of the iceberg. For example, there are now many studies that show that most people are considerably less accurate in recognizing faces of persons of a different race than they are at recognizing faces of persons of their own race. Although there is some debate over the causes of this cross-racial deficit, there is general agreement that it is real and material.

Turning to memory, there is a wealth of data indicating that a person's memory for faces never seen before fades rapidly, and while the pace of the forgetting varies considerably across individuals, there is little doubt that identifications first made in lineups or photo arrays conducted weeks after the crime in question are particularly problematic. Moreover, memory is notoriously plastic. A person who picked a photo out of a photo array a few hours after witnessing the crime will often tend, when later called to testify, to merge the crime scene and photo array memories, so that what the witness thinks are facial features she observed at the scene of the crime are actually features she had the opportunity to study, much more carefully, when viewing the photo array.

Psychological factors also influence eyewitness identifications. For example, although the Supreme Court's test suggests a strong association between accuracy and an eyewitness's "degree of certainty" at the time of the lineup or photo array, the eyewitness's assessment of how confident she is in her choice is likely to be influenced by her basic personality: some people are much more sure of their perception and memory abilities than others. ("I've always had a good memory for faces.") Thus, the Supreme Court's focus on the eyewitness's "level of certainty" at the time of initial identification appears misguided. Even if there is a relationship between eyewitness confidence and eyewitness accuracy (and the evidence for this is mixed), the association is not nearly as strong as most people (including, it would seem, Supreme Court justices) tend to think.

Moreover, in court, an eyewitness is usually asked not how confident she was when she first picked the defendant out of a lineup, but how confident she is now; and many studies have shown that, once an eyewitness has identified a particular suspect as the perpetrator, the level of her confidence will often increase over time. It is thus common for an eyewitness who said at the time of the lineup that she was "somewhat confident" in her identification of the perpetrator to later testify at trial that she is "absolutely sure" the defendant was the person she saw commit the crime.

Numerous examples of research that has revealed the complexities and limitations of eyewitness identification could be given, but let us turn to what can be done to improve the accuracy of identifications.

With respect to police procedures, it is important to distinguish between what studies have firmly established and what they simply suggest. For example, about a decade ago, some research indicated that photo arrays that were shown to an eyewitness sequentially (that is, one photo at a time) instead of simultaneously (that is, laying out all seven photos at once) led to fewer misidentifications. This, in turn, led three states (Connecticut, North Carolina, and Maryland) to pass laws requiring the sequential approach. However, some later research, as well as some statistical reanalyses of the original studies, has led some commentators to question whether the sequential approach is really better. There is also ambiguity regarding what "better" means in this context. The sequential approach may simply lead to fewer identifications period, reducing both accurate and inaccurate identifications. At present, the debate and research designed to inform it continue, suggesting that it is not yet established that one approach is superior to the other.

Nonetheless, virtually all of the most careful research done to date would support the following changes:

First, lineups and photo arrays should be blindly administered: that is, the police person administering the test should know nothing about the evidence implicating the suspect and should not know which person in the lineup or photo array is suspected of the crime (thus eliminating conscious or unconscious suggestiveness). At least nine states – Connecticut, Colorado, Illinois, Ohio, Nebraska, North Carolina, Texas, Vermont, and West Virginia – now require this.

Second, the eyewitness should be instructed that the perpetrator may or may not be in the lineup or photo array, and that the investigation will continue regardless of whether an identification is made (thus reducing any subtle pressure on the eyewitness to make an identification). A number of local law enforcement agencies have promulgated rules requiring this.

Third, the identification procedure should be videotaped in its entirety or, if this is not practical, the eyewitness's

statements should be recorded verbatim and made available to the defense. Eleven states presently require the latter.

Fourth, not just police but also prosecutors should be trained in how to avoid inadvertently influencing an eyewitness's testimony. Only a few states currently offer such training.

It must be conceded, however, that even if these and other improvements are made in police procedures affecting eyewitness identifications, the problems with eyewitnesses' own abilities to accurately perceive, retain, and recall what they saw at the time of the crime will still mean that many eyewitness identifications will to some greater or lesser degree contain inaccuracies, ranging from misidentifying the role played by someone at the scene of the crime ("I saw him fire the shot") to placing at the scene someone who was never there at all ("I know he was there because I saw him with my own eyes"). What can be done about this? Probably very little.

So far as we are aware, no one seriously suggests eliminating eyewitness testimony altogether, for many eyewitness identifications do accurately identify the culprit and get many of the details surrounding a crime correct. Without such testimony, serious crimes would go unpunished. It might be helpful, therefore, simply to inform the jury of the inherent limitations of eyewitness identifications, so that they would not let it overwhelm all other evidence or the lack of evidence. Indeed, not just judges and juries but also police and prosecutors should be trained in the limitations of eyewitness testimony and how best to evaluate its reliability.

In 2013, the Arnold Foundation asked the National Academy of Sciences to undertake a major assessment of scientific research on eyewitness performance. In response, the National Research Council (an arm of the National Academies) formed a committee to do the assessment, and it published a report in 2014 entitled *Identifying the Culprit: Assessing Eyewitness Identification*.[3] The report offers many concrete suggestions for the handling of eyewitness evidence in legal cases. Among other things, it recommends using double-blind lineups and standardized witness instructions, and it also emphasizes the need for better training of law enforcement on the potential problems of eyewitness memory. Several of the recommendations involve methods of educating the triers of fact about eyewitness memory. The information might be conveyed via expert testimony, and the authors favor giving judges the discretion to allow such expert testimony. Alternatively, information about pitfalls in eyewitness identification might be conveyed in jury instructions.

This is easier suggested than done. In a few states, notably New Jersey, judges are required to give juries detailed instructions on the many pitfalls and limitations that can threaten accurate eyewitness identification.[4] But recent studies, described below, hint that the "instruction solution" may be a form of overkill, making jurors who receive such an instruction more skeptical of all eyewitness identifications, no matter what their quality. Another alternative is to allow the parties to call experts to describe problems with eyewitness identification that might be present in the case at bar. Since expert witnesses are subject to cross-examination, their opinions might come across as less definitive than a judge's instructions. Furthermore, the other side could also call rebuttal experts. Research to date does not, however, tell us whether this use of experts would result in a better educated jury, more aware of the limitations of eyewitness identification, or simply a more confused one.

Research on the impact on jurors of court instructions and expert testimony is often problematic, since, among other difficul-

ties, the subjects are often mock jurors rather than real ones, or else they are real jurors offering post-trial self-reports that may be heavily affected by the context in which the reports are given. Indeed, one of the few studies of actual jury deliberations concluded that jurors pay greater attention to court instructions than mock-juror and post-trial self-report studies had suggested.[5] Nonetheless, relevant research has been published with respect to how specialized jury instructions and expert testimony may impact jurors' assessments of eyewitness testimony. Some such studies were precipitated by the 2011 decision of the New Jersey Supreme Court in *New Jersey v. Henderson*.[6] The underlying case involved a man named Larry Henderson who was implicated in a murder in a New Jersey apartment back in 2003. The key evidence against Henderson was the eyewitness testimony of James Womble, who had identified Henderson from a photo array. But serious problems with Womble's confident account arose later on, and when the case reached the New Jersey Supreme Court, it promulgated new rules for dealing with eyewitness testimony. Briefly, if the defendant can show any evidence of suggestive influences surrounding an eyewitness account, the court must hold a hearing in which all factors that might have a bearing on the eyewitness evidence are explored. If, after this exploration, the judge decides to admit the eyewitness evidence, the judge must provide tailored instructions that will guide jurors on how to evaluate the eyewitness evidence in the case.

These special instructions were drafted over the next year and made public in 2012. They are meant to "educate" jurors on the limits of eyewitness identification. They tell jurors, for example, that human memory does not work like a video recording that an eyewitness can replay when he wants to remember a face. The instructions also educate jurors about factors influencing eyewitness testimony that are generally accepted in the scientific literature. For example, if an identification is one in which a member of one race has identified a stranger of a different race, the instructions inform the jury that people may have a greater difficulty in accurately identifying members of races different from their own. The *Henderson* instructions have been celebrated for going further than prior instructions in providing scientific information that may aid the jury in making decisions that can have such a profound effect on someone's liberty.

But "the jury is still out" on how well these instructions achieve their intended purpose. So far, there have been only a few efforts to study what impact they might have on potential jurors who hear them. In one study, mock jurors watched a thirty-five-minute murder trial video that had either strong evidence or weak evidence of the accused's guilt, and they heard either a standard instruction or the new *Henderson* instructions.[7] A major finding was that the jurors were more than twice as likely to convict the defendant of murder when the standard instructions were used than when the *Henderson* instructions were used. However, the reduction in conviction rate when the *Henderson* instructions were used occurred regardless of whether the case was weak or strong, leading the investigators to conclude that the *Henderson* instructions do not raise doubts specific to likely inaccurate identifications, but rather induce skepticism of all eyewitness identifications.

A more recent effort examined the impact of the *Henderson* instructions in a mock jury case heard by adult community members.[8] The case was loosely based on an actual trial in which the defendant was convicted but the verdict was later overturned. Some mock jurors received the *Henderson* instructions before hearing the eyewitness testimony, while others received the *Henderson* instructions after the testimony. A major finding was that both the "before" and "after" jurors were less likely to convict the

defendant than mock jurors who heard no *Henderson* instructions. As in the previous study, however, *Henderson* instructions appeared to induce skepticism to eyewitness identifications across the board.

Neither of these studies is remotely conclusive, nor purports to be. But these admittedly preliminary studies of the impact of judicial instructions in sensitizing jurors to the limitations of eyewitness testimony at least suggest that such instructions may not adequately serve their intended function of enabling jurors to discriminate more accurately between reliable and unreliable eyewitness testimony.

What about the alternative of allowing the parties to call experts to address such issues? Although more time consuming and expensive, the use of expert testimony as a way to educate jurors can be better tailored to the case at hand than one-size-fits-all jury instructions, and does not carry the potential overweight of an instruction from the court. Numerous studies of such expert testimony have produced mixed results, however, with some studies suggesting that expert testimony does sensitize jurors to factors that affect their assessment of eyewitness testimony, while other studies show the testimony simply induces skepticism or has little impact.

For example, one study by psychologist Brian L. Cutler and colleagues concluded that an eyewitness expert does improve the ability of jurors to discriminate accurate witnesses from inaccurate ones.[9] But other studies suggest that, as in the case of specialized jury instructions, expert testimony about the limitations of eyewitness evidence simply makes jurors more skeptical.[10] So no really firm conclusions can be drawn. A recent effort attempted to compare special instructions to expert testimony more directly.[11] In this study, mock jurors watched a videotaped trial in which a defendant was charged with attempted rape. The trial lasted anywhere from forty to seventy-five minutes, depending on whether *Henderson* instructions were given, expert testimony provided, or neither or both. The authors concluded that, for the most part, neither the *Henderson* instructions nor the expert testimony did much to sensitize jurors to the quality of the eyewitness identification. What is more puzzling is that the *Henderson* instructions did not affect verdicts at all, in contrast to the substantial effect found in the earlier studies on the *Henderson* instructions described above. The authors speculate that this might be due to the specific facts of their case, or the particular eyewitness factors that they manipulated, or even the length of the experiment. In the end, however, they worry that the New Jersey Supreme Court may have been overly optimistic about the likelihood that jury instructions would improve juror evaluations of eyewitness evidence.

Despite their different results, these studies convey a similar message: there are limits on how much we can do to eliminate inaccurate eyewitness identifications. But the ultimate intractability of the problem only means that we must persevere in our efforts to mitigate it as much as possible. We owe it not only to the Kirk Bloodsworths of the world, but also to ourselves to ensure, to the best of our ability, that our criminal justice system is anchored in the truth and not simply in appearances.

ENDNOTES

[1] The Innocence Project, "Eyewitness Misidentification," https://www.innocenceproject.org/causes/eyewitness-misidentification/.

[2] *Manson v. Braithwaite*, 432 U.S. 98, 97 S. Ct. 2243, 53 L. Ed. 2d 140, 1997, U.S. LEXIS (U.S. June 16, 1997), 12.

[3] National Research Council, *Identifying the Culprit: Assessing Eyewitness Identification* (Washington, D.C.: The National Academies Press, 2014). Disclosure: coauthor Judge Rakoff was cochair of the committee that wrote the report.

[4] New Jersey Courts, "Supreme Court Releases Eyewitness Identification Criteria for Criminal Cases," press release, July 19, 2012, https://www.judiciary.state.nj.us/pressrel/2012/pr120719a.pdf.

[5] Shari Seidman Diamond, Beth Murphy, and Mary R. Rose, "The 'Kettleful of Law' in Real Jury Deliberations: Successes, Failures, and Next Steps," *Northwestern University Law Review* 106 (4) (2012): 1537–1608.

[6] *State of New Jersey v. Larry R. Henderson*, 208 N.J. 208 (2011).

[7] Athan P. Papailiou, David V. Yokum, and Christopher T. Robertson, "The Novel New Jersey Eyewitness Instruction Induces Skepticism but Not Sensitivity," *PLoS One* 10 (12) (2015): e0142695.

[8] Marlee Kind Dillon, Angela M. Jones, Amanda N. Bergold, Cora Y. T. Hui, and Steven D. Penrod, "*Henderson* Instructions: Do They Enhance Evidence Evaluation?" *Journal of Forensic Psychology Research and Practice* 17 (1) (2017).

[9] Brian L. Cutler, Hedy R. Dexter, and Steven D. Penrod, "Expert Testimony and Jury Decision Making: An Empirical Analysis," *Behavioral Sciences and the Law* 7 (2) (1989): 215–225.

[10] Michael R. Leippe, Donna Eisenstadt, Shannon M. Rauch, and Hope M. Seib, "Timing of Eyewitness Expert Testimony, Jurors' Need for Cognition and Case Strength as Determinants of Trial Verdicts," *Journal of Applied Psychology* 89 (3) (2004): 524–541.

[11] Angela M. Jones, Amanda N. Bergold, Marlee Kind Dillon, and Steven D. Penrod, "Comparing the Effectiveness of *Henderson* Instructions and Expert Testimony: Which Safeguard Improves Jurors' Evaluations of Eyewitness Evidence," *Journal of Experimental Criminology* 13 (1) (2017): 29–52.

The Uncertain Future of Forensic Science

Jennifer L. Mnookin

Abstract: Forensic science is at a crossroads. In the last two decades, often-used forms of pattern evidence, such as fingerprint, tool mark, and bite mark identification, have faced significant criticism for lacking adequate scientific validation or proven reliability. Is this the beginning of a sea change, signaling the rise of a science-based, empirically grounded approach to these forms of evidence, both in the courtroom and in the crime laboratory? Or has the increased attention produced Band-Aids rather than meaningful and lasting cures? This essay argues that the current state of forensic science reform is both "half empty" and "half full." Looking first at bite mark evidence, then at modifications in the language used by forensic scientists for their courtroom testimony, and, finally, at the creation and the elimination of the National Commission on Forensic Science, this essay argues that we have thus far seen modest and meaningful – but far from adequate or transformative – reform. Our best hope for sustained, substantial changes necessary for improving forensic science evidence within our system of justice requires the creation of another national commission or other institutional body, made up of both research scientists and other institutional stakeholders, and situated as to prevent "capture" by either forensic practitioners or advocates within our adversarial system.

JENNIFER L. MNOOKIN is Dean and David G. Price and Dallas P. Price Professor of Law at the University of California, Los Angeles, and Cofounder and Codirector of PULSE@UCLA (Program in Understanding Law, Science, and Evidence). She writes extensively about evidence law, expert and scientific evidence, and forensic science. In addition to coauthoring two treatises, *Modern Scientific Evidence: The Law and Science of Expert Testimony* (David Faigman et al., 2016), and *The New Wigmore: Expert Evidence* (with David Kaye and David E. Bernstein, 2010), her recent scholarship has appeared in *UCLA Law Review*, *Texas Law Review*, and *The Supreme Court Review*.

Forensic science evidence is at a crossroads. Over the last two decades, forensic science claims and methods have been subject to a growing chorus of academic and scientific criticism. Much of the criticism has focused on the deeply inadequate research foundations of many forms of regularly used pattern identification evidence, including latent fingerprints, tool marks, bullets, bite marks, documents, and signatures. Important reports by experts and from authoritative institutional bodies such as the National Academy of Sciences (NAS) and the President's Council of Advisors on Science and Technology (PCAST) have forcefully expressed concerns about the reliability and validity of these and other forensic science techniques like blood spatter evidence and arson determinations. This attention and criticism have expanded public awareness and spurred nontrivial reforms and meaningful institutional and research engagements.

doi:10.1162/DAED_a_00523

Still, as of yet, there has been little fundamental change in how forensic science is used in courtrooms around the country.

This present reality – a host of meaningful but mostly superficial changes alongside a still-faltering trickle of serious research – permits two radically different stories to be told about the likely future of the forensic sciences over the next decade or two. It would be possible (though just barely) to tell a "momentum" story, suggesting that we are on the cusp of an increasingly empirically based, science-driven approach to the validation and use of these influential kinds of evidence. But an equally possible – and, in my view, more realistic – story would characterize the changes made thus far as genuine, but limited and sputtering efforts at reform, unlikely to operate as gateways toward necessary substantial transformations, at least on the near horizon.

Looking at present-day forensic science is thus akin to peering at one of those well-known optical illusions like the figure on the following page, where, from one perspective, the viewer sees a young woman looking away from the observer, but then by squinting or shifting the visual vantage point, the viewer sees the image transformed into an elderly lady, eyes cast downward (see Figure 1).

In this essay, I explore these two linked but disparate ways of understanding the current state of forensic sciences and their use in the courtroom. My argument is, in essence, that neither the present nor the future of the forensic sciences can be adequately understood without taking into account both perspectives at once, somehow finding a way to see both the elegant young lady and the elderly woman at the same time. Furthermore, I suggest that unless we can create a legitimate, appropriately independent institutional body to engage meaningfully with ongoing important questions about the forensic sciences – an authoritative commission or task force institutionally located at least one step outside both the adversarial system and law enforcement–linked forensic science laboratories – the chance of sustained, thoughtful reform in the service of justice is distressingly low.

I start with some basic background information about the current state of the forensic sciences, especially forms of pattern evidence, and then look more closely at three examples that illustrate how the current state of the forensic sciences is a glass both half empty and half full. We have, on the one hand, seen genuine progress toward scientifically sounder forensic science claims; but on the other hand, that progress is dispiritingly limited. Specifically, I first look at the disparate approaches taken in two recent bite mark identification cases, a revealing contrast because bite mark comparisons are one of the least probative and most poorly validated kinds of forensic science evidence currently in regular use. I then turn to examine some modest court-imposed modifications on how forensic experts may testify; and finally, I look, very briefly, at the creation and the dismantling of the National Commission on Forensic Science (NCFS). Unfortunately, with the decision not to recommission the NCFS, we now lack any locus for a broadly conceived, authoritative panel of experts and stakeholders to convene regularly to assess the state of forensic science and recommend reforms. Unless we can reproduce some space and place for those engagements, the hope of ongoing momentum for thoughtful reform and engagement substantially dims. Although I appreciate that an ongoing effort spearheaded by the National Institute of Standards and Technology (NIST) may bring some valuable, albeit limited, near-term improvements, unless we are able to create some broader site for sustained and fair-minded engagement, outside of the control of both forensic scientists themselves and stakeholders in the adversary

Figure 1
William Ely Hill's "My Wife and My Mother-in-Law"

Source: William Ely Hill, "My Wife and My Mother-in-Law," *Puck Magazine*, 1915.

system, changes are likely to be partial, faltering, incremental, and insufficient.

To situate these issues, I offer a somewhat whirlwind overview of key concerns in the forensic pattern identification fields, a broad-strokes overview of core issues and limitations that everyone who wishes to be educated in the controversies surrounding pattern evidence ought to know.

Many feature-comparison methods have been used in court (and for investigatory purposes) for decades, or in some instances, for more than a century. Expert handwriting identification evidence, for example, has a nineteenth-century origin; latent fingerprint identification was first used in the United States in 1911, in the case *People v. Jennings*; firearms and bullet comparison evidence received attention in the Sacco and Vanzetti trial in 1920 and began to be used with increasing frequency thereafter; and bite mark identification evidence appears in some Salem witchcraft trials, but its more modern origin story traces back to *People v. Marx*, a 1975 California murder case involving a bite mark on the murder victim's nose.[1]

Forensic science pattern identification methods grew up mostly outside of universities, undisciplined by norms of academic research. Rather, these methods were primarily developed to aid criminal investigations and facilitate crime-solving. From the late 1920s and 1930s onwards, they largely gestated and developed within crime laboratories, and these laboratories were usually structured as appendages to law enforcement.[2]

Until recently, most forensic scientists had law enforcement backgrounds that typically did not include substantial formal training in science. Traditionally, many forensic analysts and crime scene investigators were sworn officers, though this is much less true today. Many forensic laboratories do now require their analysts to have an undergraduate science or forensic science degree, but even now, few forensic practitioners have Ph.D.-level training in science.[3]

Given the educational backgrounds and experience of their personnel, forensic science laboratories have, unsurprisingly, generally lacked a "research culture."[4] Until recently, to the extent that there was any university-based research in support of the forensic sciences, professorial investigators were few; those that did exist tended to be closely involved with law enforcement communities and their efforts were primarily directed toward justifying or increasing the reach of the forensic sciences rather than putting their knowledge claims to hard tests or validating the specific methods used.

Indeed, until approximately the last decade, there was remarkably little serious, methodologically sophisticated research conducted on forensic science error rates or methods. We continue to have many gaps in our knowledge about the accuracy of examiners and the real-world error rates for most pattern identification sciences.[5] Given their widespread use in the courtroom, it is rather astonishing how little peer-reviewed, high-quality evidence establishes the scientific validity of often-used methods like tool mark assessments and firearms identification. More research, but still not a great deal, exists for areas like fingerprint and document identification.

The risks posed by inadequately validated forensic science are not merely theoretical or abstract. In wrongful convictions established through DNA-based exonerations, faulty forensic science appears in roughly half of the known cases, making it the second most frequently found contributing factor (second only to erroneous eyewitness identification).[6]

For what purpose is pattern evidence used in court, and how is it presented? Forensic science fields answer questions like:

Was the defendant identified by the analyst the source of the fingerprint lifted from the crime scene? Did a bullet or cartridge casing match the defendant's gun? It is important to recognize, however, that in these and other areas, the words used to describe a forensic scientist's findings, like "match" and "identification," are fuzzy and not self-explicating. Furthermore, they can have either a strong or a weak meaning. Some fields traditionally claim a strong conclusion that individualizes pattern evidence to a specific person or source, like linking a fingerprint uniquely to the defendant's index finger. In other fields, accepted conclusions take weaker forms. When microscopic hair analysis was regularly used, for example, the field's standards prohibited the conclusion that any specific individual was the definite source of a given hair. An expert was supposed to testify only that the hairs examined shared a set of class characteristics that made it *possible* that they shared a common source, meaning that hairs found at a crime scene might have come from the defendant (while the scientist was also expected to acknowledge that other individuals' hair could also match the hairs taken from the crime scene). Nevertheless, a recent audit, jointly conducted by the FBI and the Innocence Project, found that microscopic hair identification experts regularly overstated their conclusions when testifying, frequently going well beyond what the field ostensibly permitted.[7] Thus, one concern about these kinds of evidence relates to fields making too-strong, scientifically unvalidated claims (like "individualization"); another relates to examiners who go beyond the field norms and testify to stronger conclusions than can be supported by the field.

Pattern identification determinations by experts are fundamentally – and are acknowledged by practitioners to be – *subjective*. They lack formal, validated criteria for determining a match. Rather, all the forensic feature pattern fields involve a trained examiner looking closely at the questioned item and determining, based on training, experience, and judgment, whether sufficient similarity exists to claim a match.[8] As one fingerprint expert stated nearly twenty years ago (as true now as when he wrote it), determining a match requires a certain "leap of faith" through which the expert becomes "subjectively certain."[9]

Apart from DNA identification, none of the widely used forms of pattern identification evidence currently rest upon an established statistical foundation.[10] At present, pattern identification experts across the whole range of other forensic identification fields, including firearms, tool mark, microscopic hair, latent fingerprint, and bite mark identification, as well as questioned document examination, cannot assert a quantitative, validated probability that two items of evidence match or share a common source. No fully validated models provide specified statistical probability akin to the "random match probability" used with DNA evidence. (There are, however, some reasonably promising preliminary efforts under development.)

However, at the same time, at a conceptual level, pattern identification claims *are* and must be probabilistic; we just do not yet have fully specified, validated probabilistic models.[11] Forensic examiners once commonly resisted the claim that they were engaged in a probabilistic enterprise, but there is now a growing acceptance among forensic thought leaders (if not, perhaps, the rank and file) that comparison judgments are indeed probabilistic at their core. Perhaps foreshadowing changes to come, the Department of Defense fingerprint laboratory recently became the first to require its analysts to use likelihood ratios (based on an internally developed model) in their testimony to express the strength of a match in quantitative terms, but their model has

not yet been subjected to formal validation or broad peer review.

Forensic practitioners typically do undergo proficiency tests that to some degree assess the accuracy of their judgments. Seldom, however, do these tests adequately mimic actual casework conditions. Not only are the samples included in these tests often easier to evaluate than many samples encountered in actual forensic work, but most often, these tests are not conducted "blind"; that is to say, analysts know they are being tested, which may induce them to take extra care.[12] Although it is possible to insert realistic forensic samples within the stream of casework, blinding protocols for proficiency testing are rarely used. Best practices, following the model of randomized controlled trials in medicine, in which even the person providing the sample to the analyst does not know it is a test sample, are virtually unknown.[13]

Many forensic methods raise concerns about cognitive bias. Given that most forensic laboratories are affiliated with, or under the direct control of, law enforcement, analysts may view themselves as part of the law enforcement "team." In addition, in many laboratories, forensic analysts may be privy to significant nonforensic case information and evidence, which risks inadvertently contaminating or biasing their judgments, no matter how ethical and well-intentioned an analyst may be. This problem is made more acute by the inherent subjectivity of the methods involved.[14] A number of commentators have therefore argued for the development of protocols to prevent examiners from having access to inculpatory (or exculpatory) information unless or until it is needed for their forensic analysis.[15] Some laboratories have taken steps in this direction, but it remains more the exception than the rule.

In addition to the danger of inadvertent cognitive bias, outright forensic fraud, in which bad-apple examiners knowingly lie about what they did or found, has occurred within numerous laboratories.[16] There have been cases of intentional misstating of results; exaggerations so extreme that they cannot be attributed to mere carelessness; and even "dry-labbing," where examiners report conclusions for tests they never conducted at all. The adversary system has largely failed to operate as a check on such fraud; the many known instances of fraud have only rarely been uncovered via the mechanisms of a trial, such as impeachment or on cross-examination. When a single instance of fraud is somehow discovered, triggering a review of an examiner's past actions and reports, the review often reveals many other instances of dishonesty by that examiner.

Throughout the twentieth century, defense counsel rarely challenged the admissibility or scientific reliability of pattern identification evidence. This has begun to change, albeit in a limited fashion. While still uncommon, enterprising defense counsel have now mounted a number of substantial admissibility challenges in several different fields. Though almost no effort has led to the exclusion of the forensic evidence at issue, some evidentiary challenges have resulted in modest judicially imposed restrictions or limitations on a forensic examiner's testimony, typically restricting the language the examiner is permitted to use in describing the strength and meaning of a conclusion.[17] Here, it should be noted, I am talking only about trials. We have virtually no information on the frequency with which the prosecution has shared flawed forensic science testimony with defense counsel who then used it to persuade a client to accept a plea deal.

Until recently, forensic practitioners in many of these pattern fields regularly testified in the language of total certainty rather than probability. They also frequently testified to being "100 percent confident" of their conclusions and sometimes claimed

(preposterously) that the error rate of their technique was "zero."[18] While some practitioners still testify using language of this sort, many now make somewhat less absolute claims; indeed, there is a significant movement within the forensic science community to establish norms to preclude claims of such blatantly excessive certainty.

In 2009, the National Research Council, the research arm of the National Academy of Sciences, released a blockbuster report on the forensic sciences, which confirmed many of the limitations described above and emphasized the lack of adequate scientific validation for many forensic science fields. The report asserted, for example,

> The simple reality is that the interpretation of forensic evidence is not always based on scientific studies to determine its validity. This is a serious problem. . . . There is a notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods.[19]

This report, also calling for a new federal agency to provide oversight and assessment to the forensic sciences, seemed at first like it would be hard for courts and practitioners to ignore. However, after the report was published, some prosecutors argued in briefs and in statements that the NAS report should have no bearing on admissibility decisions or on the judicial assessment of the validity or reliability of these methods, a position strongly challenged in print by one of the cochairs of the committee responsible for the report (himself a distinguished judge).[20] While many courts assessing the admissibility of pattern evidence made mention of the report, and a number of judges evinced surprise at the dearth of strong research to validate these techniques, very few saw these weaknesses as requiring any fundamental change to their reception of these forensic sciences as legitimate forms of evidence.

The report did matter: it brought significant additional attention to pattern evidence and its weaknesses, and it no doubt helped to educate defense attorneys, some judges and prosecutors, and numerous forensic analysts themselves. As important as the NAS report has been, however, the effects of the report on trial and appellate court admissibility decisions have to date been extremely modest.[21] But it did certainly change the conversation surrounding these techniques.

Due in significant part to this report and its reception, a trickle of research into the validity of forensic science methods has begun to emerge, with a variety of results depending on the study and the field. For example, in 2011 – exactly one century after it was first admitted in a U.S. court – a team of researchers released the first published study looking at fingerprint error rates.[22] This study found a small (though nonzero) false positive error rate and a more significant though still single digit (7.5 percent) false negative error rate. By contrast, the small number of studies examining error rates in bite mark identification uniformly show distressingly high – double digits or higher – levels of error.[23]

The 2009 NAS report's call for the creation of a new, full-fledged administrative agency focused on forensic science never achieved significant political traction, but in 2013, as a partnership of the Department of Justice (DOJ) and the NIST, the NCFS was born.[24] It included a significant number of academic research scientists, as well as stakeholders from the forensic science and criminal justice communities, and emerged as a respected and significant location for addressing forensic science reform. However, after three years of operation, the Department of Justice decided not to reauthorize the Commission.

In 2016, a second distinguished group of scientists – the President's Council of Advisors on Science and Technology – issued

its own major report about pattern identification sciences.[25] This report, like the NAS report seven years earlier, found a general dearth of adequate scientific studies to establish the validity of many kinds of forensic science. PCAST's report made a variety of recommendations, including some pointedly directed at judges and their admissibility determinations. PCAST asserted the fundamental need for scientific validity as a prerequisite for using scientific evidence (both as a matter of good science and good law). It further refined the concept of validity as it applies to the forensic sciences by identifying two key parts: foundational validity for any given field, and the validity of an analysis as applied in particular laboratories to particular casework.

The fundamental tenet of the PCAST report was that foundational scientific validity requires appropriate empirical studies to establish accuracy and error rates. These studies are a nonnegotiable *sine qua non* for which there simply can be no substitute.[26] In the report's words:

> Without appropriate estimates of accuracy [and error rates], an examiner's statement that two samples are similar – or even indistinguishable – is scientifically meaningless: it has no probative value, and considerable potential for prejudicial impact. Nothing – not training, personal experience, nor professional practices – can substitute for adequate empirical demonstration of accuracy.

PCAST asserted that many forms of forensic pattern-matching evidence presently lack foundational validity. PCAST found that simple DNA analysis and latent fingerprint identification do have enough appropriate testing to establish foundational validity, but that numerous other fields, including firearms and tool mark identification, bite mark identification, and more complex DNA mixture analysis, do not. The PCAST report strongly intimated that these long-used forms of evidence ought not to be admitted at trials unless or until foundational validity (and validity as applied) can be established; and that some kinds of evidence, like bite mark evidence, were likely never to be proven valid.

Given these strong conclusions, it is perhaps unsurprising that prosecutors gave the report a strikingly chilly reception.[27] Nor, unfortunately, do most courts seem inclined to take seriously PCAST's framework and admonitions, despite PCAST's high status and distinction, and notwithstanding the obligation of judges in the federal courts and in many states to play an explicit gatekeeping role in which they assess expert and scientific evidence for validity.

This broad-strokes overview of the past and present of forensic science illustrates the deep cultural divide between forensic science practice and research science. The pattern identification disciplines grew up within law enforcement, not universities; their methods are subjective and experience-based rather than objective or statistical; and forensic practitioners even today typically lack doctoral-level science training. Judgment honed by experience is the primary coin of the realm, not formal empirical study or statistical modeling. In many fields, we still lack substantial, validated information about how likely analysts are to offer inaccurate conclusions. Consider this point carefully. These forms of evidence are being used to establish guilt for serious criminal offenses, yet we lack substantial knowledge of how often the evidence presented is mistaken, overstated, imprecise, or wrong.

Furthermore, for most of their history, the fundamental epistemic legitimacy of these experience-based disciplines was almost never challenged in court. At best, the biases or competency of a testifying examiner was questioned or an effort was made on cross-examination to get an examiner to hedge a bit about the certainty of a conclu-

sion. When defense challenges to admissibility are mounted now, courts sometimes insist upon modest changes to the words used by the expert, but generally admit the evidence, notwithstanding the lack of substantial scientific testing and validation. Several distinguished interdisciplinary panels of scientists and other experts have weighed in with great concern for, but thus far only modest influence on, the lack of adequate scientific foundation and validation in these fields. Finally, we must note the real, human costs to using problematic forms of forensic evidence in court, including the danger of wrongful conviction.

This overview of key points about forensic science shows that pattern identification fields have faced modest, but not transformative, reform. Half full: the emergence of some credible research; the recognition by many practitioners that they are engaged in a probabilistic enterprise; substantial focus on these issues from two extremely prestigious, highly legitimate bodies of experts, NAS and PCAST; the creation of the NCFS and a NIST-led effort that calls on the forensic science disciplines to identify standards and approaches in need of reform. Half empty: relevant research remains limited and many fields still lack adequate validation; though probabilistic in theory, trial testimony remains grounded on experience and subjective judgment; both judges and practitioners have largely resisted the conclusions of the NAS and PCAST reports; the NCFS no longer exists.

I turn now in somewhat more detail to three specific examples that show, simultaneously, the existence of some degree of meaningful change, alongside reasons for, at best, muted optimism about further transformation, at least in the near term.

First, bite marks. On June 8, 2017, Alfred Swinton was released from prison after serving eighteen years of a sixty-year sentence for murder. His release came about because the bite mark identification that had been a lynchpin of the prosecution's case at the original trial was no longer deemed persuasive or valid by anyone involved (including the prosecutor's office and the original bite mark expert himself). Faced with this consensus, the judge vacated the original conviction, ordered a new trial, and released Swinton to house arrest while awaiting further judicial proceedings.

Once suspected of being a serial killer in Connecticut, though never charged with other crimes, Swinton's conviction was vacated not just because of the crumbling of the credibility of the bite mark evidence, but also because DNA testing showed that biological material from the bite mark (on the victim), as well as on the victim's nails, did not match Swinton's DNA. In addition, a bra in Swinton's possession had been said at trial to have belonged to the victim, but DNA retrieved from the bra did not match that of the victim. Without the DNA corroboration of Swinton's claim of innocence, perhaps the prosecutor would have stood by the state's original contention that the bite mark identification constituted substantial proof of guilt. Even with the DNA evidence and the discrediting of the bite mark evidence, the state has not yet conceded Swinton's innocence.

But the bite mark expert's disavowal of his earlier testimony pulled no punches: "I no longer believe with reasonable medical certainty – or with any degree of certainty – that the marks on [the victim] were created by Mr. Swinton's teeth, because of the recent developments in the scientific understanding of bite-mark analysis," odontologist Constantine Karazulas told the *Hartford Courant*.[28] He even called his earlier testimony "junk science" and stated that he "no longer believes that Mr. Swinton's dentition was uniquely capable of producing the bite marks I observed. . . . Indeed, many thousands of individuals could have produced those in-

juries."[29] (Interestingly, Karazulas used a novel form of computer enhancement to examine Swinton's dentition in the original trial; the issues surrounding this use of computer-enhanced images produced a lengthy, detailed Connecticut Supreme Court opinion affirming the legitimacy of the computer enhancement techniques he used, while breezing over the question of the reliability of bite mark identification in a mere footnote.)[30]

Does this case indicate a potential sea change for bite mark evidence, one of the most problematic forms of pattern evidence in current use? At the time of the original trial, the expert had called his own techniques "the new gold standard for forensic odontology" and celebrated his care and confidence in his conclusion of a match.[31] Now, instead, he offers a forceful recantation of his earlier claims. Given this about-face by the expert, coupled with DNA evidence that generally failed to link Swinton to the murder, the prosecution neither defended the legitimacy of the bite mark evidence nor opposed the defense's motion to vacate. If the prosecution decides to retry the case, they appear ready to acknowledge that bite mark identification evidence will have no legitimate role in the next go-round.

By contrast, just a couple of months earlier, in a retrial of a murder case in Pennsylvania vacated for constitutional flaws, a state court trial judge ruled bite mark evidence admissible. In motions preceding the trial, the judge even denied Paul Aaron Ross, the defendant (who was well represented with substantial involvement from the Innocence Project), the opportunity for a *Frye* hearing, in which the defendant would have argued that bite mark evidence was no longer "generally accepted" by the relevant scientific community, the standard that scientific and expert evidence must meet in Pennsylvania in order to be admissible.[32] The defendant was therefore denied the opportunity to present testimony or detailed evidence about bite mark testimony's known weaknesses. Judge Jolene Kopriva appears to have denied the hearing primarily because bite mark evidence was not a *novel* kind of proof: "The unique aspect of this case is that challenges are being brought to an existing scientific field, not a novel methodology," Kopriva said.[33] "Although the use of bite mark evidence is beginning to face challenges, it would be premature," she said, "for this court to order that the methodology is no longer generally accepted in the relevant scientific community." There is, however, a bit of a catch-22 here. If the judge will not allow a full evidentiary hearing about the validity of bite mark testimony because such evidence has long been admissible in Pennsylvania, how can she determine whether it is in fact premature to conclude that the evidence is no longer accepted in the relevant scientific community?

Of course, whether a form of evidence is deemed "generally accepted" depends in part on precisely how one defines the relevant scientific community: there is more acceptance of bite mark evidence among bite mark experts themselves than in a more broadly defined scientific community. But as we see from the Swinton case, even some bite mark experts no longer believe in the validity of the technique.[34] And there is little doubt that a great many other scientists who have examined the field have significant doubts about its validity – or, perhaps more accurately, substantial confidence that validity is lacking.[35]

To be sure, Judge Kopriva did place modest limits on the form that bite mark testimony could take. In the first trial, the expert testified that the bite marks were "very highly consistent" with the defendant's dentition. This time he would be limited to the language permitted by the bite mark expert's professional society, the American Board of Forensic Odontol-

ogists, which recommends that bite mark testimony offer one of three possible conclusions without further detail or elaboration: 1) that the person is included within the pool of possible sources; 2) that the result is inconclusive; or 3) that the person is excluded from the pool of possible biters. (At the time of writing, the defendants have requested an interlocutory appeal of the ruling to disallow a *Frye* hearing.)

More than two dozen DNA exonerations to date involve cases in which bite mark evidence played an important role at trial.[36] Additionally, unlike areas of forensic science in which the problem is a near-total lack of research (like tool mark identification), or areas in which the existing research is limited and methodologically flawed but weakly supports an inference of validity (like firearms identification), numerous bite mark studies affirmatively illustrate its significant weaknesses. (In one study, experts lacked substantial agreement on whether certain marks left on skin were bite marks at all, much less on whether a given person's dentition could have produced them.)[37]

Nonetheless, Judge Kopriva deemed it premature to take a stand, or even to permit a full admissibility hearing to make an evidence-based assessment. This invites the question: if it would be premature to exclude it now, on our current basis of knowledge, at what point, and on what research basis, could exclusion be warranted?

To be fair, if Judge Kopriva believes that a *Frye* hearing is allowable only for a novel technique, rather than a long-used one, she could feel stymied by legal doctrine and the judicial responsibility to protect settled questions from being reopened. Some cases do limit *Frye*'s ambit to the novel; but if so, how should change come about when needed? Surely *Frye* cannot mean that *any* form of expert evidence of long-standing use must be admissible forevermore, even if novel evidence of its problems, limitations, and lack of validity emerge? How ought the need for change be balanced with the legal system's norms regarding closure and precedent? Judge Kopriva's treatment of bite mark evidence illustrates the potentially awkward fit between judicial practices and the potential need to unsettle the ways we use forensic science. The law regarding the admissibility of scientific evidence may not change at the pace of science, but surely it should change when science substantially undercuts the original justifications for allowing the evidence.

We thus see a pair of encounters with bite mark evidence, in two courthouses in two different states, merely months apart, each taking a profoundly different approach to the assessment of bite mark evidence (albeit at different stages within the criminal process). In one case, we see a clear recognition of fundamental weaknesses in an especially dubious form of identification evidence. In the other case, we see the mighty power of precedent even in the face of a deeply troubling kind of proof. The Swinton case suggests change is on its way, at least with respect to bite mark evidence (perhaps in some ways too easy a target, as one of the most obviously problematic of the forensic sciences). At the same time, the Ross case suggests that even bite mark identification evidence may not be going anywhere too soon. Lest Ross be written off as a peculiar outlier, it is worth noting that not a single trial-level judge in any court has yet excluded bite mark evidence based on its lack of reliability.

One could view the dramatic difference between these two cases as a simple side effect of our federal system, which allows wide disparities in state evidence rules, as well, perhaps, as a consequence of the cases' procedurally different postures. Nonetheless, these two cases suggest rather different pathways for the future of bite mark evidence – and by extension, perhaps for the forensic sciences more generally.

In the Swinton case, we can admire the collaborative efforts to face up to the weaknesses in the bite mark testimony, including the odontologist's willingness to disavow the accuracy of his earlier testimony, as well as the prosecutor's privileging of evidentiary integrity over closure and maintaining the conviction. Building on the Swinton case, one could plausibly suggest that bite mark evidence is in its death throes. Other evidence supports this possibility too, such as the fact that the Texas forensic science commission – Texas, a state not generally seen as soft on crime – declared a moratorium on bite mark evidence until or unless a stronger research foundation could be established.[38] The glass half full, it would seem; the double image seen from its more captivating angle.

Still we cannot escape the contrast with the judge in Ross, several months earlier, refusing even to permit the evidentiary hearing that would have given the defense's concerns a full airing. The fact that she restricted the form of allowable testimony and could cite precedent for her decision is small consolation, for a person's freedom is at stake, and if bite mark evidence is fundamentally unsound, the judge's small step is akin to permitting a Ouija board's messages into evidence so long as the expert through whom the evidence is offered suggests the board is a probable rather than a certain source of truth.

In deciding as she did, Judge Kopriva is not unique. Far from it: no judge to date has rejected a prosecution offer of bite mark evidence, notwithstanding two dozen DNA exonerations in cases in which bite marks were erroneously associated with the defendant, and a growing research literature challenging the validity of bite mark identification claims. My point, however, is less to excoriate Judge Kopriva than to highlight the deep power of precedential thinking that, when rooted in outdated or pseudoscience, generates a deep – and systemically intended – legal bias in favor of maintaining the status quo, perpetuating the ongoing acceptance of questionable forensic science pattern evidence.[39] Now the glass looks half empty: if judges refuse to exclude bite mark evidence, it is difficult to imagine they will insist on stronger scientific foundations as a precondition for the admissibility of evidence involving less egregiously unreliable, but nonetheless inadequately studied, techniques. If strong, affirmative evidence that bite mark evidence is unreliable is not enough, what hope is there that courts will take the mere absence of adequate evidence seriously? So the double image switches back with a blink of an eye, the viewer as downcast and dispirited as the world-weary woman.

This pair of bite mark decisions, grounding two dramatically different narratives about what lies on the near horizon, vividly illustrates how, at present, depending on what we are looking for and from what angle we choose to see, we can find both reason for hope in and reason for hopelessness about forensic sciences' future. To understand both the present and the range of possible futures for forensic science, we must recognize that both the optimistic and the pessimistic narratives are plausible; indeed, in some meaningful sense, both are true.

Judges today are tremendously reluctant to exclude from trials long familiar forms of forensic science evidence even when, as with bite mark evidence, the scientific foundation is weak and the evidence has played an established role in numerous proven wrongful convictions. But there is a growing move – insisted upon by some judges as a precondition to admissibility, and also called for by some leaders in the forensic science community – to scale back exaggerated and overconfident assertions of knowledge and authority by forensic scientists. One author refers to this felicitously as a shift from the "dogma expert," who

asserts her findings in absolute terms, to the "transparent expert," who is more honest about the limits of her knowledge and the existence of some uncertainty in her conclusions.[40] Although this shift has begun, it remains incomplete, and the limits called for by judges and forensic leaders are often not fully enforced. Still worse, these well-intentioned efforts to rein in exaggerated forensic science claims may, in actual practice, have little meaningful effect on how forensic science evidence is actually assessed or understood by juries.

Not long ago, experts in many of the forensic science subspecialties routinely testified in the language of absolute certainty. They would, in many fields, make identifications that were both absolute and particularized: identifications made to the "exclusion of all other" fingers, or firearm cartridges, or tools, and they often claimed to be "absolutely certain" or "100 percent confident," or that their technique boasted a "zero error rate." Although such dogmatic testimonial over-claiming still occurs, it has certainly become less frequent. We now routinely see experts offering (and courts insisting upon) somewhat more epistemically modest, less hubristic claims about the established strength of the evidence. Instead of saying a fingerprint could not possibly have come from anyone in the whole world but the suspect, the fingerprint expert might now say that based on her experience, she believes that the chances of two prints sharing this much correspondence and not deriving from the same source is remote, or very unlikely, rather than absolutely impossible.[41] Moreover, this testimonial shift has spurred some examiners and labs to grow enthusiastic about a potential statistical turn, through which they could someday provide validated quantified assessments, or testify about likelihood ratios, rather than experience-based conclusions, even though not so long ago, most examiners roundly disclaimed any legitimate role for such probabilistic thinking.[42] However, as the essay by Joseph Kadane and Jonathan Koehler in this issue indicates, it is unclear whether a modest, as opposed to a major, scaling back in testimonial certainty is likely to have any effect on how probative the factfinder perceives the evidence to be.[43]

The positive, half-full narrative about this development would emphasize three points. First, by reining in unjustified overstatements that were previously commonplace, courts are not simply improving the quality of the testimony heard by the factfinder, but also increasing their own focus upon the substantive value of the offered evidence. Regulating the experts' language is modest gatekeeping perhaps, but it is a form of gatekeeping nonetheless, and therefore possibly a stepping stone to more thorough scrutiny of whether an adequate foundation undergirds an expert's claims. Second, this change in the form of testimony has spurred additional interest from the forensic science community in efforts to develop validated probabilities to ground the weight of testimony. As mentioned above, the Department of Defense announced in March 2017 that its experts would henceforth testify using quantified likelihood ratios based on an internally created statistical model.[44] Whether that model will hold up to scrutiny will be interesting to see, but the increased interest in developing such models is itself a positive turn. In addition, if forensic examiners embrace this shift to more modest language and conclusions, this focus potentially invites more careful thought about the relationship between evidentiary support and testimonial conclusion, which might in turn contribute to a broader shift in perspective, encouraging experts to see themselves less as partners to law enforcement and more as scientists.[45]

The more pessimistic, still-half-empty story would assert that these modest changes in the language used for testimony are

the forensic equivalent to rearranging the deck chairs on the Titanic. Will a factfinder – especially a lay juror primed by popular culture and shows like CSI to believe in the power of a forensic "match" – hear or weigh the evidence differently because of modestly changed language? How different is it, really, for a juror to hear that an expert believes the chances of this print coming from anyone other than the defendant is "extremely low," rather than "zero," especially if that conclusion is accompanied by an assertion like, "based on my many years of experience, I truly would not expect to see this much similarity unless the two prints came from the same person"? Or, still worse, that it would be a "practical impossibility" for them to come from different sources? The space between *impossible* and *unlikely* is real, but in this context, it may also be razor thin, especially when coupled with assertive body language and an authoritative tone of voice. In reality, the specific words used to convey the meaning of the match to a factfinder may be far less important than evidence professors or scientists might hope or think: "impossible" and "very unlikely" may, in practice, be near-fungible within the trial setting. (To be sure, this point operates in both directions: it may be that factfinders interpreted the earlier claims of "impossible" as, in fact, meaning something more akin to "very unlikely," notwithstanding the expert's stronger claim. Even so, the key point remains: modest linguistic changes in an expert's phrasing of her conclusions may have little real effect on jurors' assessment of probative value.)

The pessimistic story would also emphasize that many of the forensic science disciplines still lack adequate empirical grounds even for their weaker claims: without a valid statistical model, how does an examiner truly know that a coincidental match is as unlikely as she asserts? Furthermore, by insisting upon minor modifications to the language of testimony, judges confronting *Daubert* and *Frye* challenges may deem themselves to have taken adequate and appropriate action in response to the concerns about the validity of forensic science. Rather than acting as a spur to further engagement, modest reform in testimonial language may instead lead to judicial quiescence and complacency. In the Ross bite mark case, we see how the judge did partly constrain the testimony, but simultaneously refused the *Frye* hearing on admissibility. The optimistic story sees these minor reforms to testimony as a sign of reflective engagement by the judiciary and practitioners, and a potential stepping stone to bigger change; the pessimistic story sees it as a superficial salve that may permit the avoidance of deeper and more important cures.

One final example of the dynamics of change and its limits can be seen in the establishment of and, then, roughly three years later, the failure to renew the National Commission on Forensic Science. The Commission was a joint creation of the Department of Justice and NIST. Commissioners came from a variety of fields, including research scientists, law professors and judges, forensic science providers, law enforcement officials, prosecutors, and defense attorneys. The NCFS's mandate was, in short, "to enhance the practice and improve the reliability of forensic science."[46]

When the NCFS began, many – frankly, myself included – were dubious that a forensic science improvement effort partly centered in the Justice Department was likely to bring about significant change or meaningful improvement. Moreover, the NCFS, during its short lifespan, was not without controversy. In January 2015, Federal District Court Judge Jed Rakoff, the sole federal judge on the Commission, noisily resigned over the Justice Department's unilateral decision that issues concerning

pretrial discovery procedures for forensic evidence were beyond the Commission's legitimate scope, a decision he saw as putting "strategic advantage [for prosecutors] ahead of the truth."[47] The controversy drew media attention; the DOJ backtracked, and Judge Rakoff returned to the Commission. That controversy and the media coverage it received likely enhanced the NCFS's independence from the DOJ.

Over its three years of operation, the NCFS made a series of advisory recommendations specifically directed at the Justice Department, and offered a number of additional "views": documents that captured the Commission's collective consensus on important topics. Recommendations ranged from directing that forensic science service providers should all be appropriately accredited (which perhaps seems obvious, but accreditation has not been a uniform or consistent expectation for forensic laboratories); to stating that forensic labs should develop written policies about their documentation, reporting, and interpretation practices (this also perhaps seems obvious, but again, has not been standard practice); to an effort to enhance pretrial discovery (the subject over which Judge Rakoff nearly resigned); to a recommendation that forensic experts cease using the phrase in their testimony "to a reasonable degree of scientific [or field-specific] certainty," because the phrase, though often used in court, has no accepted scientific meaning.[48]

The NCFS recommendations were, for the most part, not transformative or field-changing. But they were thoughtful, meaningful steps in a positive direction. Moreover – and perhaps most important – the Commission had become a place where a set of thoughtful interlocutors, including academic research scientists and stakeholders in the adversarial process, could jointly consider what was needed to make forensic science fields more trustworthy, fairer, and better grounded in valid science. The significant inclusion of research scientists alongside the more typical "stakeholder" participants gave the NCFS institutional legitimacy and also meant that its proposals generally incorporated insights from scientific, legal, and forensic perspectives.

Unfortunately, the Commission proved short-lived. In April 2017, Attorney General Jeff Sessions announced that he would not renew the NCFS. I recently asked one dozen academics and scientists with an interest in forensic science to share with me their brief thoughts on the most positive and the most negative occurrence relating to forensic science in the last decade. Fully half of those I asked named the demise of the NCFS as the single worst development that had occurred in that time period.

The half-full story, then, is that the Commission existed at all, and that it managed, in a quite short amount of time, to develop both some valuable proposals and some institutional legitimacy. The half-empty story is, obviously, that it no longer exists.

As the Commissioners wrote in their final report, summarizing their activities: "The National Commission on Forensic Science has provided an essential forum … to improve the forensic sciences. … But there is still work to be done."[49]

And indeed, there is. Will that work take place? In another decade or two, will we be able to tell a story of ongoing, meaningful incremental change? Will a fair-minded observer be able to conclude that the forensic techniques we use in court have an appropriate degree of scientific validation to support their use and their conclusions? It seems reasonable to hope that the most blatantly problematic forms of forensic science, like bite mark evidence, will no longer be used. But will other kinds of forensic pattern evidence be on a surer foundation? Will testimony be presented in epistemically justified ways, and experts' claims limited to their legitimate evidentiary basis? Or will experts still testify

to near certainty without empirical validation? Will we still lack validated statistical models and robust proficiency tests?

The future of forensic science remains uncertain, but our best chance for substantial ongoing improvements rests on the creation, or re-creation, of an entity akin to the NCFS. Simply put, we need some institutional structure, some body, separate from the courts, from adversarial advocates, and from practitioners themselves, a body that includes representatives from all these arenas along with accomplished research scientists. We are simply not likely to see continued forward motion unless there is some institutional body to prompt reform, a commission or working group with both convening power and a claim to legitimacy, in which academic researchers and forensic science stakeholders can jointly assess the state of forensic science and continue to push for, and argue about, improvements.

While it is deeply unfortunate that the NCFS was not recommissioned, perhaps this institutional failure also generates an opportunity. Perhaps, just perhaps, it creates an opening for the building of a next-generation commission, one positioned wholly outside the prosecutorial and law enforcement apparatus, and yet with sufficient institutional legitimacy not to be ignored. The path to the creation of such an alternative to the NCFS is neither obvious nor easy. But if it can happen, the future of forensic science will almost certainly be far brighter, and the substance of what is used in investigations and offered in courtrooms throughout our nation will be more reliable, more trustworthy, and more scientifically valid. That vision for the future of forensic science is most certainly not assured, but it may yet be possible. There is little reason to have confidence that either the courts or the forensic science community, much less the Justice Department, will have the capacity or the will to make significant positive improvements on their own. But if we can somehow create an institutional space where scientists, lawyers, judges, and forensic leaders all work together, a collaborative space that values reason-giving, empirical research, and thoughtful engagement with evidence and its assessment, then perhaps that half-full glass may yet fill to the brim.

AUTHOR'S NOTE

Many thanks to Shari Diamond and Rick Lempert for their extremely thoughtful comments and suggestions, and to UCLA reference librarian Elyse Meyers for her excellent help. I am also grateful to the participants in the American Academy of Arts and Sciences meeting for this volume for valuable questions and ideas. I also want to honor the memory of Stephen Fienberg (1942 – 2016), University Professor of Statistics and Social Science at Carnegie Mellon University. Originally, he and I had planned to coauthor an essay for this volume. I am so sorry not to have had that opportunity, but I am also enormously grateful for the ways he used his wide-ranging curiosity, collaborative talents, and brilliance to bring insights from statistics into so many areas of public policy import, including, of course, forensic science.

ENDNOTES

[1] Sources on the history of specific pattern identification techniques include, for example, Michael J. Saks, "Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science," *Hastings Law Journal* 49 (4) (April 1998): 1069 – 1141; David L. Faigman, Edward K. Cheng, Jennifer Mnookin, et al., *Modern Scientific Evidence: The Law & Science of Expert Testimony* (Eagan, Minn.: Thomson Reuters, 2016); Simon A. Cole, *Suspect Identities:*

A History of Fingerprinting and Criminal Identification (Cambridge, Mass.: Harvard University Press, 2002); Jennifer L. Mnookin, "Scripting Expertise: The History of Handwriting Identification Evidence and the Judicial Construction of Reliability," *Virginia Law Review* 87 (8) (December 2001): 1723–1845; and *People v. Marx*, 126 Cal. Rptr. 350 (1975).

[2] J. Edgar Hoover, "The Scientific Crime Detection Laboratory," *The University of Chicago Law Review* 10 (3) (April 1943): 335–338.

[3] Exceptions include forensic odontology, in which experts have advanced degrees in dentistry, and DNA profiling, in which analysts often have graduate-level science training.

[4] Jennifer L. Mnookin, Simon A. Cole, Itiel E. Dror, et al., "The Need for a Research Culture in the Forensic Sciences," *UCLA Law Review* 58 (3) (February 2011): 725–779.

[5] The best black-box study we currently have of any forensic science discipline is Bradford T. Ulery, R. Austin Hicklin, JoAnn Buscaglia, et al., "Accuracy and Reliability of Forensic Latent Fingerprint Decisions," *Proceedings of the National Academy of Sciences* 108 (19) (May 10, 2011): 7733–7738.

[6] Brandon L. Garret and Peter J. Neufeld, "Invalid Forensic Science Testimony and Wrongful Convictions," *Virginia Law Review* 95 (1) (March 2009): 1–97; and Innocence Project, "Wrongful Convictions Involving Unvalidated or Improper Forensic Science that Were Later Overturned through DNA Testing" (New York: Innocence Project, 2016), https://www.innocenceproject.org/wp-content/uploads/2016/02/DNA_Exonerations_Forensic_Science.pdf.

[7] Spencer Hsu, "FBI Admits Flaws in Hair Analysis over Decades," *The Washington Post*, April 18, 2015, https://www.washingtonpost.com/local/crime/fbi-overstated-forensic-hair-matches-in-nearly-all-criminal-trials-for-decades/2015/04/18/39c8d8c6-e515-11e4-b510-962fcfabc310_story.html?utm_term=.100652ec7671; and Federal Bureau of Investigation, "FBI/DOJ Microscopic Hair Comparison Analysis Review," https://www.fbi.gov/services/laboratory/scientific-analysis/fbidoj-microscopic-hair-comparison-analysis-review.

[8] Simon Cole, "More than Zero: Accounting for Error in Latent Fingerprint Examination," *Journal of Criminal Law and Criminology* 95 (3) (2005): 985–1078; Melissa K. Taylor, David H. Kaye, Thomas Busey, et al., "Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach" (Gaithersburg, Md.: National Institute of Standards and Technology, 2012), https://www.nist.gov/sites/default/files/documents/oles/latent.pdf; and Jennifer L. Mnookin, "The Validity of Latent Fingerprint Identification: Confessions of a Fingerprinting Moderate," *Law, Probability and Risk* 7 (2008): 127–141.

[9] David A. Stoney, "What Made Us Ever Think We Could Individualize Using Statistics?" *Journal of the Forensic Science Society* 31 (2) (April 1991): 197–199. In some ways, this is parallel to the subjective certainty required by a juror to make the judgment to convict, but we do not call this latter judgment scientific.

[10] We have a strong statistical basis for determining the "random match probability" for DNA – that is, how likely it is that a person selected at random would match the sample in question – in cases involving either a single source for the biological material, or when there is a mixture of two people's biological material, but one of them is known (as in a rape case). Assessing probabilities becomes significantly more challenging for DNA samples involving multiple possible unknown contributors.

[11] Christophe Champod and Ian W. Evett, "Probabilistic Approach to Fingerprint Evidence," *Journal of Forensic Identification* 51 (2) (March–April 2001): 101–122; Cédric Neumann, Christophe Champod, Roberto Puch-Solis, et al., "Computation of Likelihood Ratios in Fingerprint Identification for Configurations of Any Number of Minutiae," *Journal of Forensic Sciences* 52 (1) (January 2007): 54–64; United States Department of the Army, Defense Forensic Science Center, "Information Paper" (Forest Park, La.: United States Department of the Army, 2015) [report on using "identification" in latent print technical reports]; and Heidi Eldridge, "The Shifting Landscape of Latent Print Testimony: An American Perspective," *Journal of Forensic Science and Medicine* 3 (2) (2017): 72–81.

[12] See *United States v. Llera-Plaza*, 179 F. Supp. 2d 492 (E.D. Pa. 2002) [Llera-Plaza I]; *United States v. Llera-Plaza*, 188 F. Supp. 2d 549 (E.D. Pa. 2002) [Llera-Plaza II]; Jonathan J. Koehler, "Proficiency Tests to Estimate Error Rates in the Forensic Sciences," *Law, Probability and Risk* 12 (2) (2013): 89 – 98; and Jonathan J. Koehler, "Fingerprint Error Rates and Proficiency Tests: What They Are and Why They Matter," *Hastings Law Journal* 59 (5) (2008): 1077 – 1100.

[13] Koehler, "Proficiency Tests to Estimate Error Rates" [see note 12]; and Koehler, "Fingerprint Error Rates" [see note 12].

[14] For thoughtful explorations of this problem, see D. Michael Risinger, Michael J. Saks, William C. Thompson, et al., "The *Daubert/Kumho* Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion," *California Law Review* 90 (1) (January 2002): 1 – 56.

[15] See, for example, Dan E. Krane, Simon Ford, Jason R. Gilder, et al., "Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation," *Journal of Forensic Sciences* 53 (4) (July 2008): 1006 – 1007.

[16] Mark Hansen, "Crimes in the Lab," *ABA Journal* 99 (2013): 44 – 51; Paul C. Giannelli, "The Abuse of Scientific Evidence in Criminal Cases: The Need for Independent Crime Laboratories," *Virginia Journal of Social Policy and the Law* 4 (1996): 439 – 478; and Brent J. Turvey, *Forensic Fraud* (Oxford: Academic Press, 2013).

[17] It is perhaps worth noting more broadly that the courts have shown themselves to be generally ineffective regulators of flawed forensic science. A number of kinds of forensic evidence that have now been substantially discredited (such as certain alleged indicia of arson, microscopic hair identification, and bullet-lead comparison) were regularly admitted by courts until external forces made their weaknesses impossible to ignore.

[18] See, for example, *United States v. Mitchell*, Crim. No. 960407-1 (E.D. Pa., judgment entered Feb. 2000); *United States v. Havvard*, 117 F. Supp. 2d 848 (S.D. Ind. 2000); *United States v. Hicks*, 389 F.3d 514 (5th Cir. 2004); and *United States v. Green*, 405 F. Supp. 2d 104 (D. Mass. 2005).

[19] National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (Washington, D.C.: The National Academies Press, 2009).

[20] Ibid.; and Harry T. Edwards, "The National Academy of Sciences Report on Forensic Sciences: What it Means for the Bench and Bar," *Jurimetrics* 51 (1) (Fall 2010): 1 – 15.

[21] For a general literature review looking at the NAS report's influence, see Paul C. Giannelli, "The 2009 NAS Forensic Science Report: A Literature Review," *Criminal Law Bulletin* 48 (2) (2012): 378 – 393; and "§ 29:10 Cases Citing NRC Report," "§ 32:20 Legal Responses After the NRC/NAS Report on Strengthening Forensic Sciences in the United States," and "§ 34:5 Current Developments in Caselaw" in Faigman et al., eds., *Modern Scientific Evidence* [see note 1].

[22] Ulery et al., "Accuracy and Reliability" [see note 5]. This is, in my view, a significant, thoughtful study and an important step forward for the field. However, it is important to recognize that this study was neither conducted blind nor in ecologically valid conditions: study participants were well-aware that they were participating in a critically important study of the reliability of their field. Thus, it is not clear that their degree of effort and care would mirror that of actual casework; it might be lower because no one's liberty was at stake, but it might be higher because of the critical import of this study to the field. Moreover, there was no risk of cognitive contamination because of other case-related knowledge.

[23] Michael J. Saks, Thomas Albright, Thomas L. Bohan, et al., "Forensic Bitemark Identification: Weak Foundations, Exaggerated Claims," *Journal of Law and the Biosciences* 3 (3) (December 2016): 538 – 575; and Mary A. Bush, Howard I. Cooper, and Robert B. J. Dorion, "Inquiry into the Scientific Basis for Bitemark Profiling and Arbitrary Distortion Compensation," *Journal of Forensic Sciences* 55 (4) (July 2010): 976 – 983.

[24] National Commission on Forensic Science, https://www.justice.gov/archives/ncfs.

[25] President's Council of Advisors on Science and Technology, "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods" (Washington, D.C.: President's Council of Advisors on Science and Technology, 2016), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf.

[26] Ibid.

[27] Michael A. Ramos, National District Attorneys Association to President Barack Obama (November 16, 2016), http://www.ciclt.net/ul/ndaajustice/PCAST/NDAA%20PCAST%20Response%20FINAL.pdf; Federal Bureau of Investigation, "Comments On: President's Council of Advisors on Science and Technology, Report to the President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Pattern Comparison Methods" (September 20, 2016), www.fbi.gov/file-repository/fbi-pcast-response.pdf; and Gary Fields, "White House Advisory Council Report is Critical of Forensics Used in Criminal Trials: U.S. Attorney General says Justice Department Won't Adopt Recommendations," *The Wall Street Journal*, September 20, 2016, https://www.wsj.com/articles/white-house-advisory-council-releases-report-critical-of-forensics-used-in-criminal-trials-1474394743.

[28] Dave Altimari and David Owens, "Convicted Killer, Suspect in Slayings of 4 Other Hartford Women, Could Soon Walk Free," *Hartford Courant*, March 9, 2017, http://www.courant.com/news/connecticut/hc-hartford-swinton-murder-exonerate-20170309-story.html.

[29] Ibid.

[30] *State v. Swinton*, 847 A.2d 921, n.14 (Conn. 2004).

[31] Gus Karazulas, "New Forensic Odontology Tools" (March 28, 2001), https://www.meyerinst.com/html/lucis/new-forensic-odontology-tools.pdf.

[32] *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923).

[33] Kay Stephens, "Judge Permits Bite Mark Evidence for Ross Retrial: District Attorney Can Use Testimony About Mark During Ross Murder Retrial," *Altoona Mirror*, March 9, 2017, http://www.altoonamirror.com/news/local-news/2017/03/judge-permits-bite-mark-evidence-for-ross-retrial/.

[34] Nor is Karazulas alone; Michael Bowers, a forensic dentist, has also come out strongly against bite mark evidence. See Saks et al., "Forensic Bitemark Identification" [see note 23].

[35] See, for example, those joining Saks in ibid.

[36] "§ 35:7 Erroneous Identification and Conviction," in Faigman et al., eds., *Modern Scientific Evidence* [see note 1]; and President's Council of Advisors on Science and Technology, "Forensic Science in Criminal Courts" [see note 25].

[37] See Saks et al., "Forensic Bitemark Identification" [see note 23].

[38] Joe Palazzolo, "Texas Commission Recommends Ban on Bite-Mark Evidence," *The Wall Street Journal*, February 12, 2016.

[39] To be sure, the doctrinal argument for exclusion is weaker in *Frye* states (like Pennsylvania), where the legal standard focuses on whether a novel form of evidence is "generally accepted" by the relevant scientific community. But given that *Frye* is designed to get at evidentiary integrity by assessing the views of the scientific community itself, a substantial change in views by the community ought to spur reconsideration of a previously accepted method.

[40] Eldridge, "The Shifting Landscape" [see note 11].

[41] Simon A. Cole, "Individualization is Dead, Long Live Individualization! Reforms of Reporting Practices for Fingerprint Analysis in the United States," *Law, Probability and Risk* 13 (2) (2014): 117–150. This of course begs the question of how a fingerprint expert can truly know this based on experience alone; fingerprint experts' typical tasks do not consist of searching for the prints most similar to one another yet deriving from different sources.

[42] For one example of such efforts, see Center for Statistics and Applications in Forensic Evidence, "Pattern Evidence," http://forensicstats.org/our-research/pattern-evidence/. See also Neumann et al., "Computation of Likelihood Ratios in Fingerprint Identification for Configurations of Any Number of Minutiae" [see note 11].

[43] Joseph B. Kadane and Jonathan J. Koehler, "Certainty and Uncertainty in Reporting Fingerprint Evidence," *Dædalus* 147 (4) (Fall 2018): 119 – 134.

[44] Eldridge, "The Shifting Landscape" [see note 11].

[45] Christophe Champod, "Fingerprint Identification: Advances since the 2009 National Research Council Report," *Philosophical Transactions of the Royal Society B: Biological Sciences* 370 (1674) (2015).

[46] The United States Department of Justice Archives, "National Commission on Forensic Science," https://www.justice.gov/archives/ncfs.

[47] Spencer S. Hsu, "Judge Rakoff Returns to Forensic Panel after Justice Department Backs Off Decision," *The Washington Post*, January 30, 2015. For Judge Rakoff's resignation letter, see "Full Text: Judge's Protest Resignation Letter," *The Washington Post*, January 29, 2015, https://www.washingtonpost.com/local/full-text-judges-protest-resignation-letter/2015/01/29/41659da6-a7e1-11e4-a2b2-776095f393b2_story.html?utm_term=.2e7c6951ff06.

[48] For a complete list of recommendations and views offered by the Commission, see National Commission of Forensic Science, *Reflecting Back – Looking toward the Future* (Washington, D.C.: National Institute of Standards and Technology, 2017), Appendix C, https://www.justice.gov/archives/ncfs/page/file/959356/download.

[49] Ibid.

Certainty & Uncertainty in Reporting Fingerprint Evidence

Joseph B. Kadane & Jonathan J. Koehler

Abstract: Everyone knows that fingerprint evidence can be extremely incriminating. What is less clear is whether the way that a fingerprint examiner describes that evidence influences the weight lay jurors assign to it. This essay describes an experiment testing how lay people respond to different presentations of fingerprint evidence in a hypothetical criminal case. We find that people attach more weight to the evidence when the fingerprint examiner indicates that he believes or knows that the defendant is the source of the print. When the examiner offers a weaker, but more scientifically justifiable, conclusion, the evidence is given less weight. However, people do not value the evidence any more or less when the examiner uses very strong language to indicate that the defendant is the source of the print versus weaker source identification language. We also find that cross-examination designed to highlight weaknesses in the fingerprint evidence has no impact regardless of which type of conclusion the examiner offers. We conclude by considering implications for ongoing reform efforts.

JOSEPH B. KADANE, a Fellow of the American Academy since 2010, is the Leonard J. Savage University Professor of Statistics and Social Sciences, Emeritus, at Carnegie Mellon University. His most recent book is *Pragmatics of Uncertainty* (2017).

JONATHAN J. KOEHLER is the Beatrice Kuhn Professor of Law at the Northwestern University Pritzker School of Law. He conducts research related to decision-making in the courtroom. His work has recently been published in such journals as *Journal of Criminal Law and Criminology, Jurimetrics Journal,* and *Psychology, Public Policy, and Law.*

The study of fingerprints began in a serious way with Francis Galton's book *Finger Prints* in 1892.[1] For more than a century, fingerprint results were treated by the forensic science community (and the courts) as infallible, or nearly so. In 1985, an authoritative FBI manual stated, "of all the methods of identification, fingerprinting alone has proved to be both infallible and feasible."[2] In a 2003 segment on the television news program *60 Minutes*, the head of the FBI's fingerprint unit said that the probability of error in fingerprint analysis is 0 percent, and that all analysts are and should be 100 percent certain of the identifications that they offer in court.[3] Such hyperbole is unscientific and unsustainable. As it turned out, just a few months after this program aired, the FBI was forced to admit that its top fingerprint examiners matched a print to the wrong person in the investigation of the 2004 Madrid train bombings, one of the highest profile fingerprint cases in history.[4]

doi:10.1162/DAED_a_00524

When the National Academy of Sciences (NAS) completed its comprehensive review of many of the non-DNA forensic sciences (including fingerprint evidence) in 2009, the results were shocking.[5] The NAS found that many of the most basic forensic science claims had not been validated by empirical research. In response, federal agencies and forensic science professional organizations began working in earnest to, among other things, modify the ways in which forensic scientists present evidence in court. Simple and obvious reforms such as eliminating references to "100 percent certain" identifications and "0 percent risk of error" have already taken hold. However, forensic science reformers have been largely flying blind on the question of which specific words should replace the exaggerated conclusions that forensic scientists often provide in their courtroom testimony.[6]

Fingerprint evidence has been admitted in U.S. courts as proof of identity in criminal cases for more than one hundred years. Evidence that an unknown print recovered from a crime scene (a so-called latent print) matches a known print from a suspect or other individual is rarely challenged in court and is widely regarded by the public as conclusive proof that the person whose fingerprint matched is the source of the latent print in question. This is the case even when the match is to latent prints that are partial, smudged, or otherwise of low quality, although all of these features increase both the difficulty of declaring a definite match and the likelihood of error.[7]

Fingerprint evidence has long been a powerful tool for criminal investigators and prosecutors. A latent print found on, say, a gun recovered from a shooting scene not only helps identify a person of interest for police in the early stages of an investigation, but it also may be the single most powerful proof of a defendant's guilt offered at trial. This essay looks at the use of fingerprint evidence as a tool to persuade judges and jurors at trial or, more commonly, to persuade a criminal defendant to accept a plea bargain rather than risk a seemingly certain guilty verdict. This essay, however, concerns itself only with the presentation of fingerprint evidence at trials, asking how the way a fingerprint examiner testifies about his or her results affects the weight that factfinders are likely to assign to the evidence.

In the typical case involving fingerprint evidence, a trained examiner compares one or more latent prints with various known or exemplar prints using a high-powered microscope. This process is often preceded by an automated search through a local, state, or national database. The national database includes fingerprints from approximately 120 million people. The computer search narrows the list of candidate prints and orders them so that the most likely matches appear at the top of the list. The examiner then proceeds to make pairwise comparisons between the latent and candidate prints. The end result of the pairwise comparison process (known as ACE-V)[8] is usually one of four conclusions: *identification* (the prints comes from the same source), *exclusion* (the prints come from different sources), *inconclusive*, or *unsuitable for comparison*.

Although the ACE-V process is subjective, fingerprint examiners have historically claimed that their identifications are 100 percent certain, and that there is virtually no chance that an error has occurred.[9] The precise meaning of the word "identification" may, however, vary depending on idiosyncratic definitions and usage by various parties.[10] While the Humpty-Dumpty dictum may appeal to some testifying forensic scientists ("when I use a word . . . it means just what I choose it to mean"), it is unjustified in courts of law where the interpretation of an unfamiliar or techni-

cal phrase may be the difference between freedom and incarceration for a criminal defendant.[11]

There is some basis in the broader forensic science literature for suggesting that the weight that jurors assign to fingerprint evidence will depend, in part, on the way that fingerprint examiners describe their conclusions. Psychological studies show that the way DNA results are framed affects the value that people accord to reported matches.[12] Focusing on microscopic hair results, psychologists Dawn McQuiston-Surrett and Michael Saks found that the way hair-evidence matches are described affects mock jurors' assessments of the probability that the person whose hair is said to match is the source of the unknown hair.[13] The mock jurors in their experiments assigned higher probabilities of source identification when hair evidence was described as either a "match" or "similar in all microscopic characteristics," and lower probabilities when the forensic expert estimated the number of other people in the city who would also match.

However, these studies did not find that mock jurors' judgments varied as a function of whether forensic experts went further and volunteered their own opinion that the person whose hair was said to match was the source of the hair. Similarly, an experiment conducted by legal scholar Brandon Garrett and psychologist and lawyer Gregory Mitchell in the context of short written cases that involved fingerprint evidence found that "bolstering a match with even extravagant claims about the certainty of the match and dismissals of the likelihood that someone else supplied the prints did not increase the weight given to the match."[14] Participants in their study were no more impressed with the fingerprint evidence when the fingerprint examiner said that it was "a practical impossibility" that someone other than the defendant was the source of the latent print or the examiner simply said that the defendant "matched" the latent print.[15] These investigators concluded that it really did not matter how an examiner framed a match conclusion because factfinders give "considerable weight" to fingerprint match evidence in all forms.[16]

It appears, then, that the way forensic science testimony is presented by experts will matter in some circumstances but not others. When a defendant is a member of a group of those who might be the source of an incriminating piece of evidence, testimony that fails to point out the existence of others in the set who might be the source is viewed as stronger than testimony that expressly notes that the defendant is one of a group of people who might be the source. But when a forensic scientist indicates in some fashion his or her belief that the defendant is the source of an incriminating piece of evidence, it is not clear that such add-on comments have an extra impact on factfinders.

Thanks in large part to the 2009 National Academy of Sciences report on the state of the non-DNA forensic sciences, efforts are underway to standardize and reform many forensic science practices, including the way results and conclusions are reported in court.[17] These reform efforts have thus far proceeded with little or no guidance from empirical studies. Consequently, there is a risk that proposed changes in the conclusory language used by forensic scientists will have no impact – or perhaps even an unintended impact – on police, legal decision makers, and others who rely on forensic evidence. Our essay reports on an experiment that addresses this concern. The experiment examines how people interpret and use different verbal formulations of conclusions reached by fingerprint examiners.

Six hundred jury-eligible citizens (U.S. citizens, at least eighteen years old, with

no felony convictions) were paid to answer an online questionnaire about a hypothetical legal case that included fingerprint evidence. Our mock jurors ("jurors") covered a broad and representative cross-section of the jury-eligible population in terms of education level (19.5 percent high school diploma or less, 14.5 percent graduate degrees), ethnicity (7.7 percent African American), and gender (58 percent women).

Jurors were presented with the following scenario:

> In a recent legal case, Mr. Richard Johnson was charged with robbing a convenience store. Although the perpetrator of the crime wore a hood that covered his face, Mr. Johnson became a suspect when the store owner told police that he thought the perpetrator sounded very much like one of his frequent customers, Mr. Richard Johnson. The store owner also told police that the perpetrator reached into the opened cash register with his bare hand and lifted one of the trays. When a police fingerprint examiner examined the cash register and its inside trays for fingerprints, he found 19 prints that were suitable for comparison purposes. The fingerprint examiner eliminated Mr. Johnson as a potential source of 18 of those prints. However, the fingerprint examiner was not able to eliminate Mr. Johnson as a possible contributor of one of the prints that was found on the cash register tray. At Mr. Johnson's robbery trial, the fingerprint examiner was called to testify for the prosecution. After the fingerprint examiner discussed his credentials, experience, and methods, the following exchange took place between the prosecutor (P) and the fingerprint examiner (FE):
>
> P: Now you said that you recovered 19 fingerprints from the cash register, is that correct?
>
> FE: Yes, there were 19 prints that had enough detail that I could compare them to known exemplars.
>
> P: What is a known exemplar?
>
> FE: It's a reference print – a print whose source is known. We compare the prints that we recover on objects from a crime scene with various known exemplars. So in this case, I had known exemplars from Mr. Johnson, the employees of the convenience store, and some other people. And I compared the prints that were on the cash register and cash register components with the known exemplars.
>
> P: OK, and what were your findings with respect to the prints that were on the cash register and the known exemplar print provided by the defendant in this case, Mr. Richard Johnson?
>
> FE: Well, first, I was able to exclude Mr. Johnson as a possible contributor of 18 of the 19 latent prints that were on the cash register or its various components. In other words, none of those 18 prints were made by Mr. Johnson. However, I was not able to exclude Mr. Johnson as a possible contributor of the 19th print. This 19th print was taken from the cash register tray.
>
> P: And so your bottom line conclusion is what?

At this point, jurors received one of six different single-sentence conclusions from the fingerprint examiner. In all cases, the conclusion that jurors saw was preceded by the words "My bottom line conclusion is that . . ." The six conclusory statements were as follows:

1. ". . . I cannot exclude the defendant, Mr. Johnson, as a possible contributor of that print."

2. ". . . the likelihood of observing this amount of correspondence when two impressions are made by different sources is considered extremely low."

3. ". . . in my opinion, the defendant, Mr. Johnson, is the source of that print."

4. "... in my opinion, the defendant, Mr. Johnson, is the source of that print to a reasonable degree of scientific certainty."

5. "... I was able to effect an individualization on that latent print to the defendant, Mr. Johnson."

6. "... I was able to effect an individualization on that latent print to the defendant, Mr. Johnson, to the exclusion of all other possible sources in the world."

The first conclusion ("I cannot exclude") is widely recognized as a scientifically accurate and defensible (albeit conservative) way to describe the results of a match between a known and unknown print.[18] If a known print from a suspect appears to share a common set of characteristics with an unknown print recovered from a crime scene and there are no other explainable inconsistencies, it follows as a matter of logic that an examiner would be justified in concluding that the suspect cannot be excluded as a possible contributor of the unknown print. However, a significant shortcoming of this conclusion is that it does not specify the size of the nonexcluded class of individuals.

The second conclusion reflects the language that has been recommended by the U.S. Army.[19] It is essentially a statement that the false positive error rate is "extremely low." Because this conclusion does not specify what is meant by "extremely low," it is hard for anyone to know how much weight to assign to this evidence.

The third conclusion may be defensible under the Federal Rules of Evidence, but as a purely scientific matter, it is also less defensible than the first conclusion because the examiner is making an inferential leap from evidence indicating that a suspect may be the source of a print to a personal conclusion that the suspect is, in fact, the source of that print.[20] Even if the available science gave the examiner good reason to believe that the class of people who might be the source of the print is very small, a source claim involves a degree of speculation and guesswork that extends beyond what the science can show.[21]

The fourth conclusion suffers from the same problem as the third, but is even more objectionable because it appends the impressive-sounding but scientifically meaningless phrase "to a reasonable degree of scientific certainty" to the examiner's personal opinion. The result may be inflated confidence or simply greater variability in understanding the level of confidence it is intended to convey.[22]

The fifth conclusion, which has long been favored by the FBI, goes further by replacing the "opinion" language in conclusions three and four with "individualization" language.[23] Use of this language might give the misleading impression that the science itself has unequivocally identified the source of the print.[24]

The sixth conclusion is even stronger than the fifth because it expressly states that the individualization has excluded all other possible sources in the world.

In sum, the first conclusion is the least objectionable from the standpoint of science and logic, though it is far from satisfying. The second conclusion is problematic because it does not explain what an "extremely low" chance of a coincidental match means. Conclusions three through six all involve a questionable scientific leap of faith in moving from the absence of proof that two prints come from different sources to a finding that the two prints must have come from a common source.

Returning to the experiment, after the fingerprint examiner stated his or her conclusion, the prosecutor repeated the examiner's conclusion verbatim, as many prosecutors do to ensure that jurors don't miss the examiner's conclusion. The examiner confirmed that this was indeed his conclusion.

Half of the jurors (Conditions 1 – 6) read a cross-examination that was tailored to challenge the specific conclusion used by the fingerprint examiner. For example, when the fingerprint examiner said that he was able to "effect an individualization" on that latent print to the defendant, Mr. Johnson, "to the exclusion of all other possible sources in the world" (Condition 6), the cross-examiner elicited a confession from the witness that he has not actually examined the prints of all other people in the world. Likewise, when the examiner says, "in my opinion, the defendant, Mr. Johnson, is the source of that print" (Condition 3), the cross-examiner elicits a confession from the expert witness that he is not claiming that he absolutely positively knows that the print came from Mr. Johnson's finger, to the exclusion of all other possible sources in the world. The other half of the jurors were assigned to a no–cross examination condition (Conditions 7 – 12). Table 1 summarizes the twelve conditions. Whether or not they read a cross-examination, jurors in all conditions answered the same set of questions about the case.

We asked jurors four "source" questions about the value of the fingerprint evidence for the proposition that the fingerprint belonged to the defendant, Mr. Johnson. Questions 1 – 3 and 5 used a scale that ranged from 1 (not at all) to 7 (extremely):

1. How strong would you say the fingerprint evidence is with respect to the prosecutor's claim that the fingerprint on the cash register tray belongs to Mr. Johnson (the defendant)?

2. How convincing would you say the fingerprint evidence is with respect to the prosecutor's claim that the fingerprint on the cash register tray belongs to Mr. Johnson (the defendant)?

3. How confident are you that the fingerprint on the cash register tray was left by the defendant?

4. What would you say is the probability that the fingerprint on the cash register tray belongs to the defendant? (Please provide a number between 0% and 100%.)

Next, we asked two "guilt" questions about jurors' beliefs that the defendant, Mr. Johnson, committed the robbery:

5. How confident are you that the fingerprint on the cash register tray was left by the defendant during the course of the convenience store robbery?

6. What would you say is the probability that the defendant robbed the convenience store? (Please provide a number between 0% and 100%.)

The answers participants provided to the four source and two guilt questions were all highly correlated with one another ($0.69 < r\text{'s} < 0.88$). We therefore created an aggregated "strength of evidence" index for each participant that gave equal weight to the six questions asked.

The next task was to compare the conditions with cross-examination to those without. To do so, we combined the indices for participants in Conditions 1 to 6 into an index with cross-examination. Similarly, we combined the indices for participants in Conditions 7 to 12 into an index without cross-examination. The data indicated that there was no effect for cross-examination. If anything, our subjects found the evidence without cross-examination slightly more plausible than the evidence with cross-examination, but the difference is so slight that we can ignore it. This permits us to aggregate Conditions 1 and 7, 2 and 8, and so on, giving us only six conditions (distinguished by the wording used by the fingerprint examiner). When we refer in the rest of this essay to Condition 1, for example, what we mean is the aggregation of Conditions 1 and 7 in Table 1; the same is true of all Conditions 1 – 6.

Table 1
Twelve Conditions

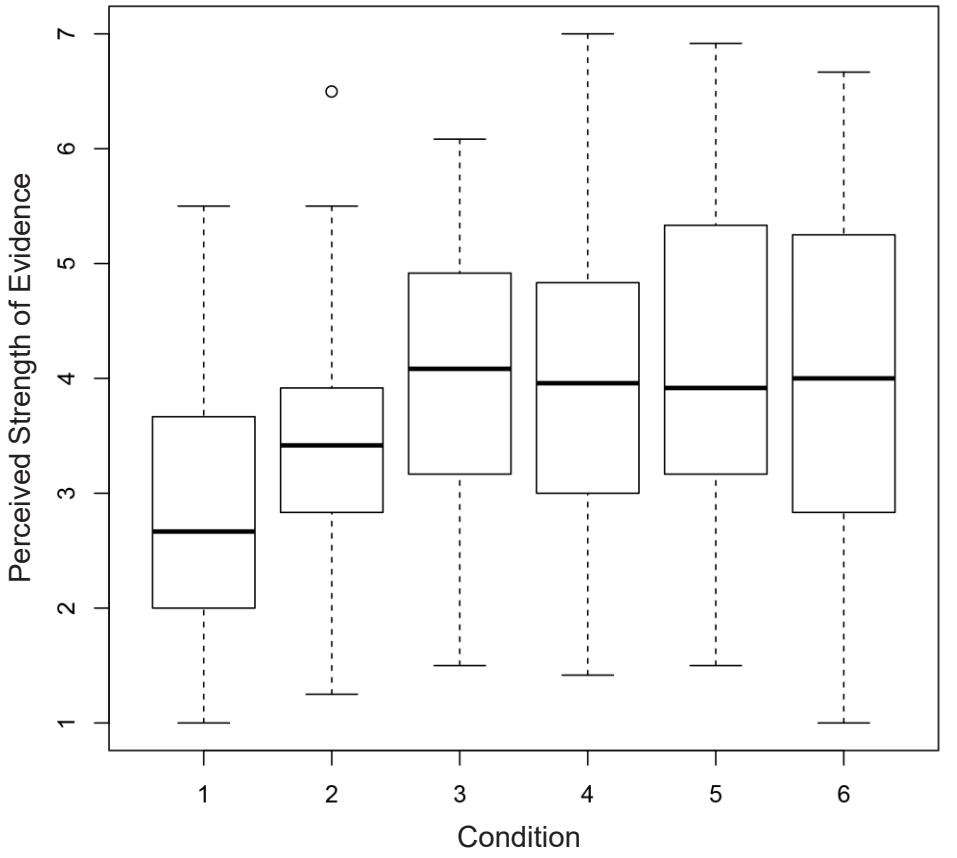
	Condition	
Expert Testimony	**Cross-Examination**	**No Cross-Examination**
Cannot exclude Mr. Johnson	1	7
The likelihood of observing this amount of correspondence when two impressions are made by different sources is considered extremely low	2	8
Mr. Johnson is the source	3	9
Mr. Johnson is the source to a reasonable degree of scientific certainty	4	10
I effected an individualization on that print to Mr. Johnson	5	11
I effected an individualization on that print to Mr. Johnson to the exclusion of all possible other sources in the world	6	12

Whether another style of cross-examination would show a larger effect is not addressed by our data.

Our primary analysis focuses on how differences in the language that the fingerprint examiner used to describe the match evidence affected subjects' judgments about the strength of the evidence. We do this by comparing the evidence strength index scores across the six fingerprint examiner conditions. In Figure 1, arrayed along the horizontal axis are the conditions, and the vertical axis is the index of strength of evidence. Figure 1 shows that the language used by the fingerprint examiner in Condition 1 ("cannot exclude") was the least impactful way of reporting the fingerprint evidence, followed by Condition 2 ("the likelihood of observing this amount of correspondence . . . is considered extremely low"). The language used to describe fingerprint evidence in Conditions 3 – 6 was more impactful than that of Conditions 1 and 2, and differed little by condition. To the extent our results generalize to actual trials, we see the importance of how forensic scientists present their testimony, and the need to ensure that the language a forensic scientist uses fairly reflects the evidentiary implications of the reported evidence.

The language used by the fingerprint examiner in the six conditions was designed to vary in the certitude with which the examiner provided his conclusion. For example, an examiner who says that he has "effected an individualization on that print to Mr. Johnson to the exclusion of all possible other sources in the world" (Condition 6) appears to be expressing much greater certainty in his conclusion than an examiner who simply says that Mr. Johnson cannot be excluded as a possible contributor of the

Figure 1
Perceived Strength of Evidence by Condition

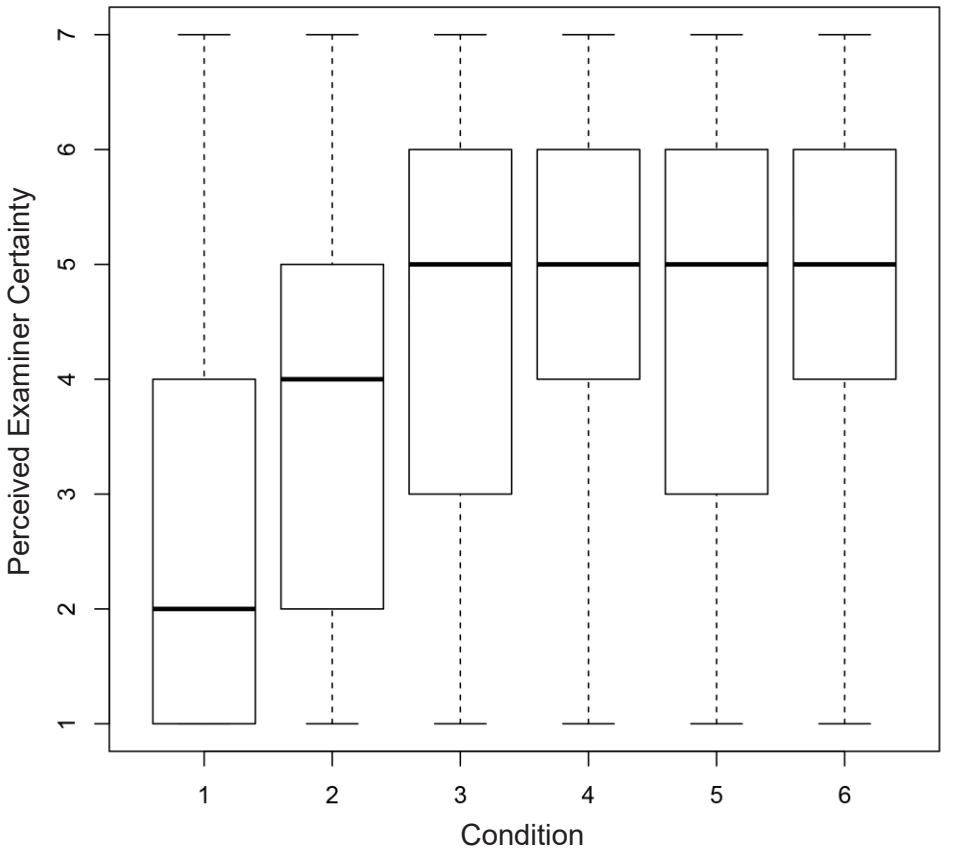


Note: Condition 1 = cannot exclude; Condition 2 = extremely low likelihood of such correspondence by different sources; Condition 3 = source of the print; Condition 4 = source of the print to a reasonable degree of scientific certainty; Condition 5 = individualization; Condition 6 = individualization to the exclusion of all other possible sources in the world. The rectangular boxes above each condition capture the interquartile range, and the bold horizontal line within each box shows the median for that condition.

same print (Condition 1). We checked this assumption by asking jurors the following question: "How certain was the expert that the defendant was the source of the fingerprint on the cash register tray?" The distribution of jurors' answers across the six conditions is plotted in Figure 2. Here the vertical axis is a scale of certainty (1 = low, 7 = high). As expected, the data show that people believe that the fingerprint examiner is least certain in Condition 1, followed by Condition 2. Further, our jurors believed that the examiner was more certain of his conclusion in Conditions 3 – 6 than in Conditions 1 and 2. It is notable that the medians for Conditions 3 – 6 are identical.

We also asked our participants whether the uncertainty expressed by the fingerprint examiner mattered: "How much does it matter in a case like this whether the ex-

Figure 2
Perceived Examiner Certainty by Condition

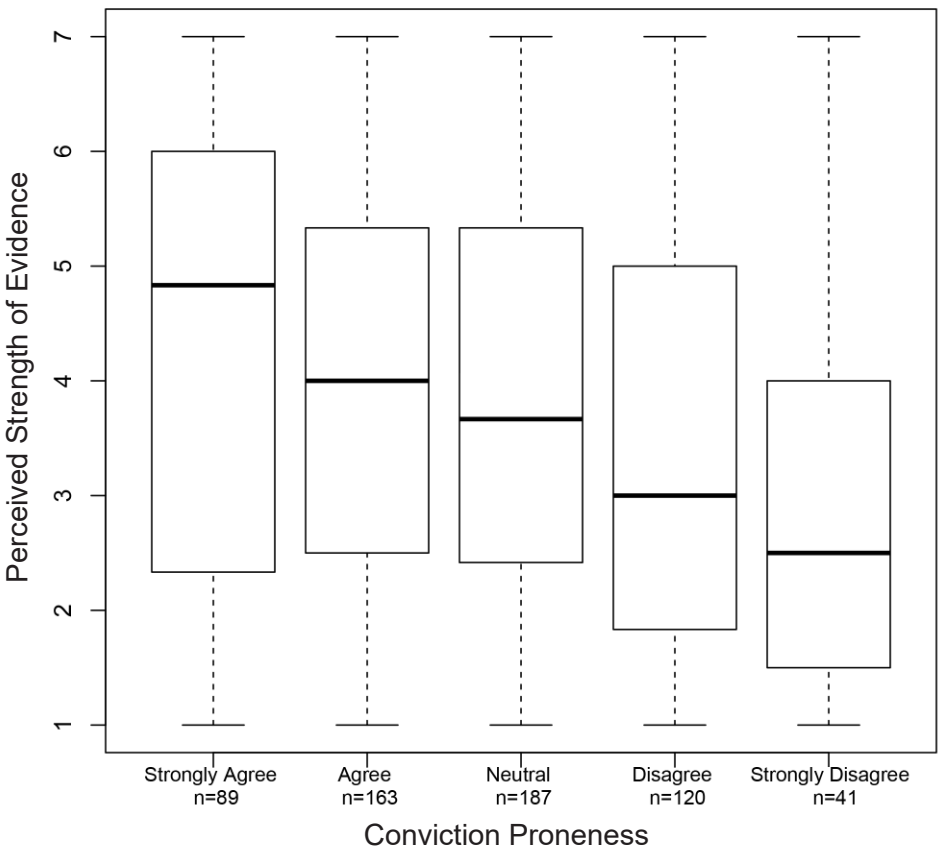


pert witness is certain about his conclusions (rather than expressing uncertainty)?" The degree of certainty expressed by the fingerprint examiner mattered a great deal to our jurors in all conditions (median ratings of 6 or 7 out of 7).

In addition to seeking judgments about the weight the fingerprint evidence deserved, we sought demographic and opinion information from our respondents. We found that men, African Americans, and those with graduate degrees are somewhat more skeptical of fingerprint evidence than others. Jury service and law enforcement service (self or relative), political leanings (liberal or conservative), and frequency of watching CSI or similar television shows had no effect on indexed responses. However, we did find a strong relationship between index scores and responses to the item below:

> Our criminal justice system should be less concerned about protecting the rights of the people charged with crimes and more concerned about convicting the guilty. (Please select only one)
>
> Strongly agree
> Agree
> Neutral
> Disagree
> Strongly disagree.

Figure 3
Perceived Strength of Evidence by Conviction Proneness



Note: Participants who marked Strongly Agree or Agree to the conviction proneness question indicated a relatively strong concern for convicting the guilty, and participants who marked Disagree or Strongly Disagree indicated a relatively strong concern for protecting defendants' rights.

Figure 3 shows that respondents who thought we should be more concerned about convicting the guilty (as reflected by agreement or strong agreement with the statement above) tended to assign more weight to the fingerprint evidence across conditions. It is not surprising that this should be so. Perhaps it is further evidence that "we see things not as they are, but as we are."

To summarize, participants in our study attached more weight to the fingerprint evidence in the four conditions (3 – 6) in which the examiner indicated in some manner that he or she believes or knows that the defendant is the source of the print than when the examiner offered a weaker, but more scientifically justifiable, conclusion (Conditions 1 and 2). The phrases commonly used to bolster source opinion and individualization claims ("to a reasonable degree of scientific certainty" and "to the exclusion of all other people in the world," respectively) had no apprecia-

ble effect on the judgments of our jurors. A simple cross-examination that was tailored to highlight weaknesses in the fingerprint evidence in each condition likewise had no impact regardless of which type of conclusion the examiner offered. Gender, race/ethnicity, education, political leaning, jury service, and law enforcement service (their own or that of a relative) produced only minor effects, though we did find that people who indicated that our criminal justice system should be more concerned with convicting the guilty tended to assign greater weight to the fingerprint evidence.

Fingerprint analysts may find that a latent print and a known print share certain characteristics. However, at the present time, they have no scientific way to estimate the number of people in a given population whose fingerprints are likely to share those characteristics.[25] In this respect, fingerprint analysis differs from DNA analysis because only the latter has systematically documented the frequency of the relevant characteristics among various populations. Consequently, there is insufficient scientific justification for a claim that the person whose fingerprint matches that of a latent print recovered from a crime scene must be the source of that print. For this reason, we believe the source and individualization statements in some of our conditions overstate the strength of the evidence.

On June 3, 2016, the Department of Justice proposed "uniform language for testimony and reports for the forensic latent print discipline."[26] This proposal included approval for a finding of identification, but barred "to the absolute exclusion of all others" and "a zero error rate or … infallible." Our results show that the proposed limitations are unlikely to affect how lay persons, such as judges and jurors, understand latent print testimony. However, we did find that when the identification language is abandoned in favor of the weaker, but more scientifically justifiable, "cannot be excluded" conclusion, people attached less weight to the fingerprint evidence. If future researchers are able to identify the frequency with which various print features arise in the population, then perhaps the cannot-be-excluded conclusion could be modified to include an estimate of the number of people who could be the source of the latent print in question. If that group is sufficiently small, it seems likely that people will attach more weight to fingerprint evidence that is presented with the further empirically justified information attached.

The President's Council of Advisors on Science and Technology (PCAST) reported in 2016 that

> Based largely on two recent appropriately designed black-box studies, PCAST finds that latent fingerprint analysis is a foundationally valid subjective methodology – albeit with a false positive rate that is substantial and is likely to be higher than expected by many jurors based on longstanding claims about the infallibility of fingerprint analysis. Conclusions of a proposed identification may be scientifically valid, provided that they are accompanied by accurate information about limitations on the reliability of the conclusion – specifically, that (1) only two properly designed studies of the foundational validity and accuracy of latent fingerprint analysis have been conducted, (2) these studies found false positive rates that could be as high as 1 error in 306 cases in one study and 1 error in 18 cases in the other, and (3) because the examiners were aware they were being tested, the actual false positive rate in casework may be higher.[27]

The two studies referred to in the PCAST report come from Noblis researcher Bradford T. Ulery and colleagues in 2011 and 2012.[28] We are less impressed with these

studies than was PCAST for several reasons. First, the subjects were volunteers who knew they were being tested. Second, the studies were paid for by the FBI (an interested party) and some of the authors worked for the FBI. Third, the proportion of judgments of identification and exclusion varied widely, suggesting that some examiners were very cautious, perhaps more cautious than they would be in casework. This reduces the credibility of the false positive and false negative rates found in these studies. Nonetheless, such studies are useful for comparing groups of fingerprint examiners and for comparing the difficulty of different types of fingerprint assessment tasks.

Although our empirical study – which obviously was not designed to measure what the Ulery studies measured – does not share these shortcomings, our results should likewise be interpreted with caution. It is a single study, conducted online with individual participants who had no opportunity to test their reactions by comparing them with those of others, and the precise wording of our stimuli and questions may have influenced the answers provided.[29] Further, because we used just one scenario and a single forensic technique (fingerprinting), it is difficult to say how well our results generalize either to other fingerprint scenarios or other forensic science methods.

Having said that, our results appear to reinforce and extend the observation by McQuiston-Surrett and Saks, and Garrett and Mitchell that lay people may not be sensitive to distinctions between stronger and weaker conclusions that an expert draws about forensic matching evidence once the expert has declared a match or words to that effect.[30] In those studies, the judgment made by mock jurors did not vary as a function of whether the forensic expert provided an opinion about whether matching hairs came from the same person (McQuiston-Surrett and Saks) or whether matching prints were described as not possibly belonging to anyone other than the defendant (Garrett and Mitchell). Likewise, our mock jurors did not draw distinctions among different "source" claims, including those that were designed to impress upon jurors that no one other than the defendant could be the source of the print. That is, once the examiner in our study offered his opinion that the defendant was the source of the print, it made no difference to our jurors whether that source claim was stated as a source opinion, a source opinion bolstered by a reference to "a reasonable degree of scientific certainty," or some form of "individualization" conclusion. Presumably, then, people process these descriptions heuristically, and reason that the expert is simply telling them: "It's the defendant's fingerprint: period." However, when the expert in our study offered a weaker, and more scientifically justifiable conclusion – one that left open the possibility that there are others besides the defendant who may be the source (see Conditions 1 and 2) – our jurors assigned less weight to the fingerprint evidence.

If the pattern of results we observed holds true across domains, then reform efforts that focus not on barring source conclusions or statements of identification but solely on eliminating the purely bolstering features of forensic match reports – features such as "to a reasonable degree of scientific certainty" or "to the exclusion of all other possible sources in the world" – may be ineffective. Although these claims may be unscientific and unhelpful, banning such language from the courtroom may have little practical effect on how jurors think about and use the forensic evidence they hear. If courts will not allow source attribution statements unless and until scientists can offer compelling scientific data that support such statements, then source attribution statements in any form should be prohibited at trial. In contrast, moving forensic experts toward more conservative, scientifically defensible claims such as "the defen-

dant cannot be excluded as a possible contributor of the print," could represent an important change.[31]

Jurors in our study also assigned relatively less weight to fingerprint evidence when the Army's language was used ("the likelihood that we would observe this degree of correspondence when two impressions are made by different sources is considered extremely low").[32] Here, as well, the perception of lower probative value probably reflects an understanding that the examiner's statement does not preclude the reasonable possibility that people other than the defendant might have prints that matched the latents in the case.

The fact that cross-examination on the shortcomings of the forensic conclusion had no impact on our jurors is disheartening. But this result is not entirely surprising. Koehler's results from a 2011 shoeprint study are similar.[33] He found that defense attorney cross-examination of a shoeprint expert had no effect on his mock jurors, even when that cross-examination revealed important risks that were ignored by the match statistic provided.[34] But it is important to remember that our cross-examination was only cursory and in print. It is possible that a well-tailored live cross-examination would be more effective.

Meaningful reform related to the way fingerprint evidence is reported should bring with it an acknowledgment that the available science does not enable examiners to prove that only one person could be the source of an unknown print. Source conclusions, including those that imply a kind of objective certitude (such as "individualization") are little more than the subjective, untested opinions of examiners. In the words of the respected forensic scientist David Stoney, such conclusions represent "a leap of faith . . . a jump, an extrapolation, based on the observation of highly variable traits."[35]

Squaring scientific accuracy with public understanding of the value of forensic science evidence will require a greater focus on empirical research, both to explore further the scientific basis of fingerprint analysis and to identify ways to convey accurately what the science has to offer and its associated uncertainties. We see no place in this endeavor for individualizations and untested source opinions.

AUTHORS' NOTE

This essay was presented at the *Dædalus* authors' conference in Cambridge, Massachusetts, in July 2017, and we gratefully acknowledge the many helpful comments made by the conference participants. We also thank Shari Diamond, Brandon Garrett, and Rick Lempert for their helpful comments on previous drafts.

The material presented here is based upon work partially supported under Award No. 70NANB 15 H176 from the U.S. Department of Commerce, National Institute of Standards and Technology. Any opinions, findings, or recommendations expressed in this material do not necessarily reflect the views of the National Institute of Standards and Technology, nor the Center for Statistics and Applications in Forensic Evidence. This research was also supported, in part, by the Northwestern Pritzker School of Law Faculty Research Program.

ENDNOTES

[1] Francis Galton, *Finger Prints* (New York: MacMillan, 1892).

[2] Federal Bureau of Investigation, *The Science of Fingerprints: Classification and Uses* (Washington, D.C.: U.S. Government Printing Office, 1985), iv.

[3] *60 Minutes*, "Fingerprints," January 5, 2003.

[4] FBI National Press Office, "Statement on Brandon Mayfield Case," May 24, 2004, https://archives.fbi.gov/archives/news/pressrel/press-releases/statement-on-brandon-mayfield-case; and Office of the Inspector General, Oversight and Review Division, *A Review of the FBI's Handling of the Brandon Mayfield Case* (Washington, D.C.: U.S. Department of Justice, 2006).

[5] National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (Washington, D.C.: National Academies Press, 2009).

[6] Spencer S. Hsu, "FBI Admits Flaws in Hair Analysis Over Decades," *The Washington Post*, April 19, 2015; "The Justice Department and FBI have formally acknowledged that nearly every examiner in an elite FBI forensic unit gave flawed testimony in almost all trials in which they offered evidence against criminal defendants over more than a two-decade period before 2000."

[7] Although latent print examiners generally will not call an identification in cases where the print quality is extremely poor, it is worth noting that when identifications *are* called on a low-quality print, they are commonly called at the same 100 percent confidence level that is used for identifications that involve very high-quality latent prints.

[8] ACE-V stands for Analysis, Comparison, Evaluation and Verification. President's Council of Advisors on Science and Technology, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* (Washington D.C.: Executive Office of the President, 2016), 9.

[9] Bradford T. Ulery, R. Austin Hicklin, JoAnn Buscaglia, and Maria Antonia Roberts, "Accuracy and Reliability of Forensic Latent Fingerprint Decisions," *Proceedings of the National Academy of Sciences* 108 (19) (2004): 7733 – 7738. Apparently, these claims of extreme accuracy have made their mark on the public. A recent study found that the median estimate mock jurors provided for the false positive error rate associated with fingerprint analysis is 1 in 5.5 million. See Jonathan J. Koehler, "Intuitive Error Rate Estimates for the Forensic Sciences," *Jurimetrics Journal* 57 (2017): 153 – 168.

[10] National Institute of Standards and Technology and National Institute of Justice, *Latent Print Examination and Human Factors: Improving the Practice Through a Systems Approach* (Gaithersburg, Md.: National Institute of Standards and Technology, 2012), 13 – 18, http://www.nist.gov/manuscript-publication-search.cfm?pub_id=910745.

[11] Lewis Carroll, *Through the Looking Glass* (Minneapolis: Lerner Publishing Group, 2002), 247. See Dawn McQuiston-Surrett and Michael J. Saks, "Communicating Opinion Evidence in the Forensic Identification Sciences: Accuracy and Impact," *Hastings Law Journal* 59 (2008): 1159 – 1189. "Forensic expert witnesses cannot simply adopt a term, define for themselves what they wish it to mean, and expect judges and juries to understand what they mean by it"; ibid., 1163.

[12] Jonathan J. Koehler, "When are People Persuaded by DNA Match Statistics?" *Law and Human Behavior* 25 (5) (2001): 493 – 513; Jonathan J. Koehler, "The Psychology of Numbers in the Courtroom: How to Make DNA Match Statistics Seem Impressive or Insufficient," *Southern California Law Review* 74 (2001): 1275 – 1306; Jonathan J. Koehler and Laura Macchi, "Thinking About Low-Probability Events: An Exemplar Cuing Theory," *Psychological Science* 15 (8) (2004): 540 – 546; Jonathan J. Koehler, "Linguistic Confusion in Court: Evidence From the Forensic Sciences," *Journal of Law and Policy* 21 (2) (2013): 515 – 539, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2227645; Dale A. Nance and Scott B. Morris, "Juror Understanding of DNA Evidence: An Empirical Assessment of Presentation Formats for Trace Evidence with a Relatively Small Random-Match Probability," *The Journal of Legal Studies* 34 (2) (2005): 395 – 444; Jason Schklar and Shari S. Diamond, "Juror Reactions to DNA Evidence: Errors and Expectancies," *Law and Human Behavior* 23 (1999): 159 – 184; and William C. Thompson and Erin J. Newman, "Lay Understanding of Forensic Statistics: Evaluation of Random Match Probabilities, Likelihood Ratios, and Verbal Equivalents," *Law & Human Behavior* 39 (4) (2015): 332 – 349.

[13] McQuiston-Surrett and Saks, "Communicating Opinion Evidence in Forensic Identification Sciences" [see note 11]; and Dawn McQuiston-Surrett and Michael J. Saks, "The Testimony of Forensic Identification Science: What Expert Witnesses Say and What Factfinders Hear," *Law and Human Behavior* 33 (5) (2009): 436 – 453.

[14] Brandon Garrett and Gregory Mitchell, "How Jurors Evaluate Fingerprint Evidence: The Relative Importance of Match Language, Method Information, and Error Acknowledgment," *Journal of Empirical Legal Studies* 10 (3) (2013): 484 – 511, 497.

[15] Ibid., 489, 495.

[16] Ibid., 497.

[17] See National Research Council, *Strengthening Forensic Science in the United States* [see note 5]. The recent decision by current Attorney General Jeffrey Sessions not to renew the National Commission on Forensic Science (NCFS) – a thirty-member advisory panel convened during the Obama administration "to enhance the practice and improve the reliability of forensic science" – introduces a large dose of uncertainty into the forensic science reform movement. See U.S. Department of Justice Archives, "National Commission on Forensic Science," https://www.justice.gov/ncfs (accessed April 28, 2017); and Erin E. Murphy, "Sessions is Wrong to Take Science Out of Forensic Science," *The New York Times*, April 11, 2017. The NCFS had been the leading force in the forensic reform movement over the past few years.

[18] Expert Working Group on Human Factors in Latent Print Analysis, *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach* (Washington, D.C.: U.S. Department of Commerce, National Institute of Standards and Technology, 2012), 129.

[19] Defense Forensic Science Center, *The Use of the Term "Identification" in Latent Print Technical Reports* (U.S. Army Criminal Investigation Command, 2015). At the time of writing, the U.S. Army appears to be in the process of replacing this language with language provided by statistical interpretation software known as FRStat. See, for example, Heidi Eldridge, "The Shifting Landscape of Latent Print Testimony: An American Perspective," *Journal of Forensic Science and Medicine* 3 (2) (2017): 72 – 81, 80. See also Arizona Forensic Science Academy, Forensic Science Lecture Series Fall 2017 Workshop Announcement, https://www.azag.gov/sites/default/files/sites/all/docs/azfsac/Announcement%20Fall%202017%20Workshop_LPC.pdf.

[20] See Federal Rules of Evidence, Rule 702, Testimony by Expert Witnesses; and David A. Stoney, "What Made Us Ever Think We Could Individualize Using Statistics?" *Journal of the Forensic Science Society* 31 (2) (1991): 197 – 199.

[21] Michael J. Saks and Jonathan J. Koehler, "The Individualization Fallacy in Forensic Science Evidence," *Vanderbilt Law Review* 61 (1) (2008): 199 – 219.

[22] U.S. Department of Justice, "Proposed Uniform Language for Testimony and Reports for the Forensic Latent Print Discipline" (Washington, D.C.: U.S. Department of Justice, 2016).

[23] Simon A. Cole, "Who Speaks for Science? A Response to the National Academy of Sciences Report on Forensic Science," *Law, Probability and Risk* 9 (1) (2010): 25 – 46 [see "effect individualizations" statement by FBI Examiner Melissa Gische on page 36]; and Peter E. Peterson, Cherise B. Dreyfus, Melissa R. Gische, et al., "Latent Prints: A Perspective on the State of the Science," *Forensic Science Communications* 11 (4) (2009), https://archives.fbi.gov/archives/about-us/lab/forensic-science-communications/fsc/oct2009/review.

[24] Jonathan J. Koehler and Michael J. Saks, "Individualization Claims in Forensic Science: Still Unwarranted," *Brooklyn Law Review* 75 (4) (2010): 1187 – 1208.

[25] As Rick Lempert pointed out to us (personal communication, June 26, 2018), even if there were data that showed that no two people had the same fingerprint, one could not conclude that no two people could leave the same print because latent prints are commonly partial or smudged. And because latent prints are incomplete or smudged in various ways, it is not clear how we might go about estimating the frequency of those prints.

[26] U.S. Department of Justice, "Uniform Language for Testimony and Reports Initial Draft for Public Comment," https://www.justice.gov/archives/dag/forensic-science (accessed June 15, 2017).

[27] President's Council of Advisors on Science and Technology, *Forensic Science in Criminal Courts*, 149 [see note 8].

[28] Ulery et al., "Accuracy and Reliability of Forensic Latent Fingerprint Decisions" [see note 9]; and Bradford T. Ulery, R. Austin Hicklin, JoAnn Buscaglia, and Maria Antonia Roberts, "Repeatability and Reproducibility of Decisions by Latent Fingerprint Examiners," *PLOS One* 7 (3) (2012): e32800.

[29] Koehler, "Linguistic Confusion in Court" [see note 12].

[30] McQuiston-Surrett and Saks, "Communicating Opinion Evidence in Forensic Identification Sciences" [see note 11]; McQuiston-Surrett and Saks, "The Testimony of Forensic Identification Science" [see note 13]; and Garrett and Mitchell, "How Jurors Evaluate Fingerprint Evidence" [see note 14].

[31] We favor supplementing such conclusions with data, derived from rigorous empirical studies, that will help legal decision makers gauge the probative value of the reportedly matching items. Once such data are collected, it may well turn out that many fingerprint matches are as probative with respect to the source question as are DNA matches (holding aside the issue of the risk of human error that may differ across methods and personnel).

[32] Defense Forensic Science Center, *The Use of the Term "Identification" in Latent Print Technical Reports* [see note 19].

[33] Jonathan J. Koehler, "If the Shoe Fits, They Might Acquit: The Value of Shoeprint Testimony," *Journal of Empirical Legal Studies* 8 (2011): 21 – 48.

[34] Ibid., 39.

[35] Stoney, "What Made Us Ever Think We Could Individualize Using Statistics?" 198 [see note 20].

Alternatives to Traditional Adversary Methods of Presenting Scientific Expertise in the Legal System

Nancy Gertner & Joseph Sanders

Abstract: The twin goals of any litigation are to arrive at a correct outcome and provide the parties with a sense that they were treated justly, even if they do not prevail. Adversarial proceedings are often perceived to be superior to inquisitorial proceedings with respect to the second goal but inferior with respect to the first. This is especially the case when proceedings involve expert testimony. In this essay, we discuss several relatively minor changes to typical adversarial processes that offer the potential of improving trial accuracy without disrupting the overall structure of adversarial proceedings. These changes include 1) alterations to the organization of the trial, including concurrent expert testimony; 2) alterations to the role of the jury, including taking notes, asking questions, and receiving written expert reports; and 3) formal expert witness codes of conduct designed to better arm experts to resist the adversarial pressures that lead to biased testimony.

NANCY GERTNER is a former Senior United States District Judge of the United States District Court for the District of Massachusetts.

JOSEPH SANDERS is the A. A. White Professor of Law at the University of Houston Law Center.

(Complete author biographies appear at the end of the essay.)

Before considering alternatives to traditional adversarial methods of presenting scientific evidence in court, we should ask why such alternatives are important in the first instance. What objectives are we seeking to achieve? In what respect do the existing methods fall short of those objectives? Federal Rule of Evidence 102, a rule that sets forth the purposes of the rules of evidence, is a useful place to begin this inquiry.

> Rule 102. Purpose
>
> These rules should be construed so as to administer every proceeding fairly, eliminate unjustifiable expense and delay, and promote the development of evidence law, to the end of ascertaining the truth and securing a just determination.[1]

The rule sets forth several goals that inform the topic of our essay. Somewhat rearranged, they are:

doi:10.1162/DAED_a_00525

1) adopt "fair" procedures that produce just determinations; 2) ascertain the truth; and 3) control the time and expense of litigation. In this essay, we focus our attention on the first two goals, but we note the many occasions when time and expense issues undermine their implementation.

As Rule 102 suggests, ascertaining the truth is not the sole objective of the rules controlling the admissibility of evidence. Nor is it the sole goal of the legal system. Procedural and evidentiary rules should be fair and should secure a just determination of the issues. "Fair" and "just determination": these terms do not define themselves.

A particularly useful approach to understanding what just determinations entail is the procedural justice perspectives of social psychologists John Thibaut, E. Allan Lind, and Tom R. Tyler, often collaborating with legal scholar Laurens Walker.[2] Their research tells us that the decision maker's perceived neutrality, the level of respect the decision maker gives the parties, the amount of voice and control the parties have over the litigation, and the trust parties have in the decision maker's motive to be fair all contribute to perceptions of procedural justice.[3] Moreover, these perceptions are remarkably robust across demographic groups and societies.[4] Although each of these factors matters when disputes are litigated, the most important components of perceived procedural fairness appear to be voice and control.[5]

Perhaps the most significant finding of the procedural justice literature is that perceptions that procedures are fair and just matter because when losers perceive that procedures have been fair and just, they are more likely to accept unfavorable outcomes.[6] Procedural fairness is a heuristic that signals how a person is likely to respond to an outcome.[7]

Rule 102 also calls for ascertaining the truth. This lofty goal is best understood as a desire to achieve factual accuracy. Factual accuracy is more easily defined than justice, but it is at least as difficult to measure. Moreover, unlike perceptions of procedural justice, which are to a large extent subjective, factual accuracy is ostensibly about objective truth, albeit a truth that is not easy to ascertain.

Individuals, both in and out of the legal system, may have opinions about how litigation structures affect outcome accuracy. These opinions may be more or less grounded in observations of the outcomes of actual cases, but objectively there are many disputes in which it is impossible to know for certain the truth against which to measure the accuracy of a case outcome.[8] Even when cases turn on expert testimony, ascertaining outcome accuracy with respect to the subject matter of the expert testimony may sometimes – not always – be possible.

We review the social science research on how these two goals map onto dispute settlement styles, specifically adversarial and inquisitorial forms of litigation. This research suggests that while adversarial processes score relatively high on procedural justice measures, they perform less well in terms of accuracy. We then discuss a number of alternatives to traditional adversary methods that may improve outcome accuracy, without sacrificing procedural fairness.

Scholars frequently divide legal systems into two categories: those that adopt adversarial processes and those that adopt inquisitorial processes.[9] However, no actual dispute resolution system perfectly exemplifies either an inquisitorial or an adversarial system. Rather, it is better to think of these as what Max Weber called "ideal types."[10] Several of the proposed changes discussed below move U.S. procedures away from a pure adversarial system, but not all the way to an inquisitorial system.

Among the important attributes of an inquisitorial system is investigation and

control of fact-finding by a neutral decision maker; among the important attributes of an adversarial system is party control of fact-finding and the evidence presented to the decision maker.[11] Legal scholar Oscar Chase, for example, included the following three adversarial system components in his essay on what he called American procedural exceptionalism: the use of party-controlled pretrial investigation, the relatively passive role of the judge at the trial or hearing, and the method of obtaining and using expert opinions (that is, parties hire and prepare experts).[12] Collectively, these components reinforce the idea that expert knowledge is a partisan resource.

How do the objectives of procedural justice and factual accuracy map onto these dispute resolution processes? We review the evidence on perceptions of justice and accuracy, followed by research that attempts to measure the objective effects of the adversarial process on accuracy.

Thibaut and Walker asked U.S. students to imagine different mechanisms for resolving hypothetical conflicts. The alternatives included mediation, investigation and resolution by a neutral decision maker, and arguments presented by advocates for each side to a third-party decision maker. The third alternative is most similar to adversarial procedures. A substantial majority of students declared this alternative to be the most just.[13]

Numerous other studies have reproduced this result.[14] Of particular interest is a recent study by litigation scholar Justin Sevier.[15] He presented his subjects with a pair of vignettes, the most relevant of which involved an expert witness in a products liability drug case. In that vignette, the subjects read a dispute concerning whether a drug manufacturer's product caused the plaintiff's illness. Some subjects were told that the case would be decided using an adversarial process and some that it would be decided using an inquisitorial process. The adversarial procedure manipulations track Chase's components of adversarialism discussed above: party-controlled pretrial investigation, passive judiciary, and party selection and employment of experts. As in earlier research, the subjects rated the adversarial procedure higher on the justice dimension, but they rated the inquisitorial procedure as more likely to produce accurate and correct outcomes.

Two things are worth mentioning with respect to this and other studies. Although the studies produce consistent results about the perceived strengths and weaknesses of adversarial processes, the effects are, as Sevier notes, modest. It is not the case that respondents perceive adversarial processes as wholly devoid of accuracy or inquisitorial processes as wholly devoid of procedural fairness. Based on these results, it would be a mistake to believe that, even at the level of perceptions, the situation is either-or: a choice between justice and accuracy.

Second, while these perceptions come from individuals who have relatively little experience with the day-to-day workings of the adversarial process, similar judgments are shared by those who do.[16] A well-known article by legal scholar John Langbein argued that German inquisitorial civil procedure was superior to adversary procedures in terms of fact-finding.[17] Other academics have expressed similar views.[18]

Not surprisingly, U.S. judges and practitioners for the most part have praised the adversarial system – in whole or in part – as a vehicle for uncovering the truth, although, to be sure, there are dissenters.[19] Three of the most notable judicial critics of adversarial procedures as a way to ascertain the truth are Federal Appellate Judge Jerome Frank, FDR's attorney general Thurman Arnold, and U.S. Federal Judge Marvin Frankel.[20] Describing adversarial proceedings as a "fight model," they reject the

argument that this approach is a particularly useful way to assess the truth. Consider, for example, this statement from Arnold:

> Bitter partisanship in opposite directions is supposed to bring out the truth. Of course no rational human being would apply such a theory to his own affairs. . . . [M]utual exaggeration of opposing claims violate(s) the whole theory of rational, scientific investigation. Yet in spite of this most obvious fact, the ordinary teacher of law will insist (1) that combat makes for clarity, (2) that heated arguments bring out the truth, and (3) that anyone who doesn't believe this is a loose thinker.[21]

There is a fair amount of empirical evidence indicating that, with respect to expert testimony, a full-blown adversarial model presents challenges to accurate fact-finding. This is not the place for a full discussion of this literature.[22] However, two types of studies are worth mentioning: those suggesting the process produces biased witnesses and those suggesting that jurors experience some difficulty with expert testimony.

Even though many judges support the adversarial system as a method of trying cases, surveys indicate that many also agree with at least one of the critiques of Frank, Arnold, and Frankel: expert witness bias. The most frequently articulated problem is that party experts abandon objectivity and become advocates for the side that hired them.[23] Several studies support this concern, indicating that simply being part of an adversarial process results in a skewed presentation of facts and opinion.[24]

In one study by business administration scholar Max Bazerman and colleagues, the subjects were professional auditors. Each of the 139 subjects was given five ambiguous auditing vignettes and asked to judge the accounting. Half the subjects were asked to suppose that they had been hired by the company they were auditing and half were asked to suppose they were hired by a different company doing business with the company in question. They were then asked to state whether or not the firm's financial reports complied with generally accepted accounting principles (GAAP). For all five vignettes, the auditors were on average 30 percent more likely to find that the accounting behind the company's financial report complied with GAAP if they were playing the role of auditor for the firm.[25]

Because parties select witnesses, bias may also arise from the selection process itself. As a number of scholars have noted, lawyers commonly consult multiple experts until they find some who reflect their position, even if this position is well out of the mainstream of expert opinion on the issue.[26] Neither of these concerns means that the experts are intentionally lying. Experts perceive others to be more biased than they themselves are, sometimes called a bias blind-spot.[27] As long as experts are chosen by parties, we should anticipate expert bias. But even within a system of adversarial selection, steps can be taken to reduce partisanship.

Although experts may present skewed, and even incomplete, evidence, this does not mean that factfinders necessarily reach erroneous decisions. In the United States, where jury trials are a central part of civil and criminal litigation, the degree to which expert bias affects outcome accuracy is typically framed in terms of jury competence. One suggested source of error in civil cases – criminal cases reflect different biases as we describe below – is that jurors uncritically accept expert testimony, a fear sometimes expressed in appellate opinions.[28] However, considerable evidence suggests that, at least when both sides present expert testimony, if anything, the opposite is true. Where there is a battle of experts, jurors are very skeptical of party experts.[29]

Although skepticism about the testimony of expert witnesses may be a good starting point for juror efforts to sort out the truth, this does not ensure accuracy if skepticism simply causes jurors to disregard expert testimony. Perhaps the clearest statement of this problem is to be found in one juror's comment offered in an asbestos case studied by law and psychology researchers Jane Goodman, Edith Greene, and Elizabeth Loftus: "The expert testimony was not a real factor in our decision, except in the very backhanded sense that it lent medical credence to any result."[30]

Skepticism seems likely to be increased by cross-examinations that cast doubt on the qualifications and biases of expert witnesses and their testimony. But in the laboratory, even "strong" cross-examinations that focus on the specific weaknesses of an expert's testimony appear to be of limited efficacy in persuading jurors to reject testimony based on faulty science.[31]

A more useful way to think of juror responses to expert testimony is to consider the distinction between central and peripheral processing.[32] In central or systematic processing, people examine the content of a communication to assess its validity.[33] In peripheral or heuristic processing, people do not attend to the quality and validity of arguments. Rather, they take shortcuts to determine the value of a persuasive attempt. People rely on factors such as the number of arguments made (rather than their quality) or attributes of the communicator such as credentials or attractiveness.

The principal question for those interested in how factfinders respond to expert witnesses concerns the degree to which the factfinders employ central processing and the degree to which they employ peripheral processing.[34] Studies have found that people centrally process information, which is to say they then engage in a high degree of cognitive processing of the message content, when they are knowledgeable about the topic, when the topic is relevant to their concerns, when they are motivated, and when the information is comprehensible to them.

In many ways, a trial is a nearly ideal setting for inviting central processing. By virtue of the role into which jurors are cast, the formality of proceedings, and the obligation that at the end of the day they will make critical decisions, jurors should be motivated to understand and reach correct conclusions based on the substantive information presented, and studies of actual jurors find almost uniformly that jurors take their task quite seriously. On the other hand, if expert evidence is complex or hard to understand to the point of being almost inaccessible, jurors are less likely to centrally process it. The other evidence in the case, which they can comprehend, will be centrally processed, while hard to comprehend expert evidence may be more peripherally processed, if not entirely ignored.

Not all peripheral processing is a bad thing. For example, a wise juror would assign some weight to the credentials and experience of experts. However, some research suggests that in assessing experts, laypeople focus more on the background and experience of witnesses than on the empirical support for a proposition.[35] While jurors may also benefit from a sense of where the weight of expert opinion lies on a particular issue, an adversarial trial often makes this difficult. Adversarial methods of selecting and cross-examining witnesses, combined with limits on the number of experts (one or two for each side on any given issue) may well make all experts appear qualified and all contested questions ones about which reasonable experts can disagree.[36]

Other peripheral cues, apart from credentials and experience, are less diagnostic of expert knowledge. For litigators, the best strategy is always to make the expert content as clear as possible and the peripheral

cues as favorable as possible, even those peripheral cues, such as an expert witness's personality and appearance, that are not at all diagnostic of the merits of an expert's arguments.[37] But if one side clearly has the worst of the science on its side, an equally plausible strategy is to make the substantive issue as confused and complex as possible, pushing jurors toward more peripheral processing or even toward discounting the expert testimony entirely.

One caveat to the preceding discussion is in order. Expert testimony has different resonance in criminal than in civil cases. It is rare that criminal cases become battles between equally competent prosecution and defense experts. Too often the adversaries are not of equal stature and lack equal resources. Rules of criminal procedure are unlikely to level the playing field. The federal criminal rules, for example, permit discovery of the basis for an expert's testimony, but do not permit the kind of robust fact discovery that occurs in civil cases.[38] Appointed counsel have a difficult time getting the state to pay for their experts at a level that would attract the best in their profession; the government has no such problem. As a result, rather than coequal dueling experts, the more typical pattern is that the government offers expert witnesses and the defendant seeks to exclude or challenge them on *Daubert* grounds, but does not offer its own expert.[39] This is especially the case with respect to the forensic sciences, a number of which, like bite mark evidence and hair analysis, as David Baltimore, David S. Tatel, and Anne-Marie Mazza discuss in this volume, have been discredited following DNA exonerations.[40]

In some criminal trials, the judicial role is simultaneously passive and active. It is passive with regard to the admission of government experts, and active in the exclusion of defense experts.[41] The process produces biased government experts, defense counsel without the tools to challenge them, and jurors who have no problem finding government experts credible, notwithstanding the heightened burden of proof in criminal cases.

The imbalance is exacerbated by the fact that the vast majority of criminal cases end in guilty pleas. The difference in resources and in judicial attitudes toward forensic evidence offered by the government as compared to the defense plays out in differences in bargaining power, which may not be immediately apparent. Since the prosecutor reasonably predicts that the trial court will be more likely to accept a government expert, however flawed or subject to meaningful challenges, and unlikely to accept a similarly situated defense witness, the government comes to the bargaining process with far more expert tools at its disposal (for example, "they identified your bite marks on the victim, so even though you say you are not guilty, the court is likely to admit the evidence and the jury to convict."). Indeed, too often defense counsel will not be prepared to offer defense expert witnesses of any caliber because she predicts the court will reject them and does not want to waste scarce resources on the effort. Government expert witnesses then become bargaining chips, their deficiencies unexamined.

Where expert testimony on traditional forensic sciences is offered in criminal cases, an imperfect adversarial process compromises both procedural fairness and truth. Ironically, in this context, most reforms seek to make that imperfect adversary system more adversarial, with discovery reforms, expert compensation, and defense training.

In sum, the perception of laypeople, academics, and judges is that adversarial processes are preferable in terms of procedural justice but may produce less accurate results. What empirical data we have suggest

that these perceptions have merit, especially with respect to expert testimony in civil cases. Given this state of affairs, some have suggested wholesale changes to the ways expert testimony is introduced into civil cases and changes in the training and makeup of factfinders.[42] Some of these changes are discussed in the contributions of Daniel Rubinfeld and Joe Cecil, and Valerie Hans and Michael Saks in this volume.[43] Criminal cases, as we have noted, raise entirely different challenges and may require even more fundamental reforms.

Our goal is more modest and hopefully achievable. We explore changes in the presentation and reception of expert testimony that fall largely within the existing adversarial structure but that promise some modest improvement in outcome accuracy. These changes have two goals: to counter the biasing effect that party witness selection has on expert testimony and to reduce the perceived complexity of expert testimony so as to facilitate factfinder central processing. We discuss several proposed changes and note how they work toward these two goals.

Since the order of presentation – at least in civil cases – is within the court's discretion, a judge could make adjustments for the purpose of making expert presentations more understandable to jurors, which could minimize the bias that comes from testimony shaped by party examination.[44] For example, the court could allow each side to present experts with one following the other, enabling a dialogue between them.[45] Likewise, the court could require that all experts testify at the conclusion of the merits trial, building on all the facts in the record.

To be sure, whether these measures affect procedural fairness depends on other orders of the court; for example, the extent to which the parties have been given discovery of expert opinions, and the timing of that discovery. And, as with concurrent testimony described below, the procedure may not be appropriate for all cases. It works best when the experts at least share some common premises, are speaking similar methodological languages, and so on.

Moreover, since the gross order of the trial is typically keyed to the burden of proof – the party bearing the burden of proof goes first and last – absent a rule change, diverging from that order would require the consent of the parties. A defendant in a civil case, intending to challenge the sufficiency of the evidence at the conclusion of the plaintiff's case, may well object to offering the defense expert for examination until he or she has had an opportunity to make such a motion.[46] Nonetheless, it would still be possible to adjust the order of expert testimony so that opposing experts, if they did not appear back-to-back, testified in closer proximity than they currently do when the parties are left to their own devices.

Even if a rule altering the order in which expert testimony is presented would aid in the evaluation of scientific evidence, such a rule would be particularly problematic in criminal cases. There is no constitutional obligation to present a defense at all. Defense counsel may well oppose presenting its expert testimony in proximity to the government expert, after the court has ruled that the government's case was sufficient to avoid a defense motion to dismiss. The government would resist an offer to delay its scientific evidence until the defense presented its experts because without the scientific evidence, the court might have to dismiss the case. These same considerations mean that rearranging the order of expert testimony by agreement is also less likely in criminal cases.

The second proposed change would follow courts in Australia, which have adopted a practice known as "hot-tubbing" or "concurrent evidence." The parties' experts, although selected in the usual way with the

usual biases, present their opinions in an unusual way after all of the lay evidence has been admitted. They are sworn as witnesses, but rather than sit with their "side," or even in the witness box, they are seated at the same table. Prior to their joint appearance, they will have filed a summary of their positions in light of the evidence. One expert then gives an oral narrative presentation, followed by questioning conducted not by counsel, but by the opposing expert; then the procedure is reversed. Each expert gives a final summary – again as a narrative, presumably accounting for more of the opposing expert's concerns – followed by conventional cross-examination by the parties.[47] In effect, as one scholar has described it, this is a procedure that has been "grafted onto" existing adversarial processes, with conventional cross-examination following a jointly presented narrative.[48] While in the usual trial, experts prepare through formal written responses and depositions led by counsel, when there is hot-tubbing, the experts are given an opportunity to speak, to comment on the evidence offered by their counterpart, and to ask questions. Moreover, they are required to meet pretrial, preferably without lawyers, on at least one occasion.[49]

In Australia, all players in the litigation have to consent to this deviation from usual adversarial procedure. The judge must also consent because her role is greatly transformed. She is necessarily a critical and active player: she suggests topics for discussion and may well pose questions to the experts.[50] The parties must agree to changes in the presentation of evidence, either concurrent evidence (the hot-tubbing approach) or consecutive.[51] Likewise, the experts who appear must agree to be bound by a code of conduct for expert witnesses, one that makes it clear that their paramount duty is to the court and not to any party in the proceedings.[52] We discuss this code of conduct below.

Hot-tubbing works best when the expert testimonies concern the same evidence or related evidence. Not only is it available in civil cases, it is recommended in Australian criminal trials as well. However, in Australia, the procedure takes place in a procedural setting that guarantees far more discovery than in the usual U.S. criminal case. Moreover, the same kinds of constitutional and pragmatic issues that are likely to limit the consecutive presentation of opposing evidence in U.S. criminal cases may also affect the likelihood that hot-tubbing becomes an accepted way of presenting expert evidence in the United States.

Nevertheless, the approach deserves more study if only because Australian judges are generally pleased with this procedure. They believe it enhances communication, comprehension, and decision-making, and reduces partisanship through the "physical removal of an expert from his party's camp to the proximity of a (usually) respected colleague."[53] The method is even praised for increasing judicial economy.[54] Experts reportedly are happy with the approach because their testimony is less skewed by conventional cross-examination. They are allowed to expand on their opinions, not limited by the traditional – and misleading – "yes or no" responses that advocates in the United States are typically allowed to insist on.[55] Presumably, the process also enhances the decision maker's ability to assess the testimony.[56]

But there are additional, perhaps even more fundamental, problems. Hot-tubbing is not easily adapted to trials with juries. Civil juries are not currently used in the Australian federal court system. While it is ironic that concerns about juror comprehension is the reason a hot-tub procedure has been proposed for U.S. trials, the jury substantially complicates its application. The parties may well be concerned about testimony by narrative, rather than testimony controlled through question and an-

swers. Admissibility issues are more complex before the jury, particularly with respect to the factual predicates of experts' opinions. And the judge's more active role in hot-tubbing could well raise concerns about showing an apparent bias for one side or the other before lay decision makers.

Even if methods are developed to use hot-tubbing with juries, criminal cases raise the additional problems of finding and paying for experts of equivalent firepower to be placed in the hot tub.[57] When substantial differences in resources between the parties result in large differences in the quality of the parties' experts, the procedure could well be problematic.

Nor is hot-tubbing appropriate for all kinds of cases, let alone all kinds of experts. As one scholar explains: "the kind of compromises that can be negotiated between town planners or geographers in relation to the size of a building or the uses of land, for example, might not be appropriate in a professional negligence proceeding or between forensic scientists in criminal matters."[58]

The verdict on hot-tubbing or consecutive experts then is: it depends – on the case, criminal and civil, on other more fundamental procedural reforms, and on judicial training.

Many of the reforms that address the adversarial system's skewing of expert testimony propose an enhanced role for the judge. Indeed, as described above, the desirability of such a change is the premise at the heart of the Supreme Court's decision in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*[59] Yet, although the Supreme Court's decision in *Daubert* was expected to initiate a new era of the judge as a more active gatekeeper in evaluating forensic evidence, it has not done so, especially in criminal cases. The 2009 National Research Council (NRC) report *Strengthening Forensic Science in the United States*, for example, criticized lawyers and judges for being "utterly ineffective in addressing the problem" of the limitations of forensic science disciplines.[60] Notwithstanding the admonitions of the NRC report and the more recent report of the President's Council of Advisors on Science and Technology (PCAST), little has changed.[61] Judges have continued to be passive with respect to the admission of expert testimony on trace evidence in criminal cases, even when scientific weaknesses have been documented, while being relatively more aggressive in screening experts in civil cases.[62]

Given this track record, it is not at all clear that U.S. judges will willingly involve themselves in meaningful supervision of a procedure like hot-tubbing. If their *Daubert* review in criminal cases has been perfunctory, as is often the case, judges will be confronting experts who are testifying based on widely divergent premises; for example, the government bite mark expert announcing that a bite mark is a match, the defense expert calling the conclusion "junk science" and challenging the field's fundamental validity. On the civil side, where judges are already more engaged, at least at the federal level, because of the resources of the parties and the judge's role in case management under Rule 16 of Federal Rules of Civil Procedure, for example – more active judicial participation may well be possible. But again a caveat: unless a judge engages with the issues and carefully examines the expert reports, judicial participation runs the risk of being more about efficiency than substance, more about case management than fairness.

"Management" sounds neutral, even salutary. One can "manage" a trial to ensure that experts are speaking to the critical issues, addressing methodological problems, and presenting their testimony clearly to a lay decision maker. Or one can "manage" a trial to ensure that it is efficient – that it ends on time and does not interfere with the court's docket – without engaging sub-

stantively with the issues. Management to promote efficiency is important, but it may or may not enhance truth seeking or procedural fairness. The critical question is the *content* of the judicial intervention, not the *fact* of it.

To change the patterns described here with respect to criminal cases may require a sea change on a number of fronts: judicial training, the allocation of resources, even the criminal rules.

One way of enhancing juror comprehension is to allow jurors to ask questions and take notes during the trial. Giving the jurors a more active role, having them behave more like adult learners, will facilitate their paying attention, their engaging with the issues, and their understanding the science they are offered. Written jury instructions, organized like a handbook, with a table of contents to accompany the jurors to the jury room, could enhance their ability to understand what the expert evidence they have heard implies for the legal issues.[63] Courtroom technology is also critical. Jurors, particularly younger jurors, are used to getting information through screens.[64] Some courts are experimenting with technology in the jury room: white boards and computers (disconnected from the Internet), for example. In addition, jurors typically perform better when they have an opportunity to take notes.[65] In one experiment, jurors permitted to use checklists and keep notebooks achieved better understanding of scientific issues than jurors who did not use these tools.[66] A more radical proposal would be to give jurors the written reports of the experts. Psychologist Lynne Forster Lee and colleagues found that jurors who received written reports summarizing an expert's testimony before the expert testified showed enhanced understanding of the testimony.[67] Still other reforms are discussed in the essay by Hans and Saks in this volume.[68]

Permitting jurors to ask questions requires careful administration because juror questions can tread on areas that are inadmissible, irrelevant, or prejudicial. The judge should inform the jurors at the beginning of the trial that it may not be possible for the court and parties to answer all of their questions. Judicial control should be exercised by requiring questions to be submitted by jurors in written form, previewing questions in advance, and sharing and discussing them with counsel outside the hearing of the jury. Studies of juror questions for witnesses during civil trials indicate that the vast majority are relevant and permissible. In particular, juror questions for expert witnesses indicate that jurors take the opportunity to submit questions for experts in an effort to understand and evaluate the content of the testimony. Similar studies have not been conducted on criminal trials, where juror questions about whether the defendant has a criminal record may be more challenging to deal with than the ubiquitous questions about insurance in civil cases.[69]

Many proposals have been advanced that alter the role of the expert by reducing allegiance to a party. Some propose greater use of court appointed experts under Federal Rule of Evidence 706. Other proposals suggest more modest alterations in the unbridled discretion of parties to select their own experts.[70] Rubinfeld and Cecil discuss some of these options in their essay in this volume.[71]

A still more modest proposal is a more complete code of ethics designed to arm an expert to resist adversarial pressures. Numerous professional organizations have adopted provisions in their codes of professional ethics dealing with expert witnesses. Although the codes differ in several ways, they share one common theme. Insofar as possible, the expert is urged to act as an objective, disinterested partici-

pant whose job is to educate the court by providing specialized knowledge on issues relevant to the case.[72] In effect, the expert's role is that of an educator.

To our knowledge, no court in the United States has adopted such a provision. In contrast, the Civil Procedure Rules of New South Wales, Australia, includes a basic statement of the expert's "general duty to the court":

> An expert witness has an overriding duty to assist the court impartially on matters relevant to the expert witness's area of expertise.
>
> An expert witness's paramount duty is to the court and not to any party to the proceedings (including the person retaining the expert witness).
>
> An expert witness is not an advocate for a party.

The code of conduct also includes a section on "[e]xperts' reports" that, inter alia, contains the following provisions:

> If an expert witness who prepares an expert's report believes that it may be incomplete or inaccurate without some qualification, the qualification must be stated in the report.
>
> If an expert witness considers that his or her opinion is not a concluded opinion because of insufficient research or insufficient data or for any other reason, this must be stated when the opinion is expressed.

The code of conduct must be provided to each expert and the expert's testimony or the written report cannot be entered into the case unless the expert acknowledges the receipt of the code of conduct and agrees to be bound by it.

Setting aside quibbles over the specific language in this code, the underlying purpose is salutary. It emphasizes that the expert's duty is to the court and reinforces the ideal of a disinterested educator role for experts.[73] Even if the code of conduct is entirely hortatory without any sanctions for its violation, it could effect some improvement.[74]

Why should we wish to change the ways in which expert testimony is presented in U.S. courts? There are several possible answers. One could argue that alternatives could produce more efficient trials, or could reduce the very high cost of litigating cases involving expert testimony. In this essay, we focus on still another reason. The alternatives presented above attempt to balance the twin goals of producing more accurate outcomes and retaining procedures that are perceived to be fair. For example, hot-tubbing puts some limits on the parties' control over their testimony, but it still permits parties to choose their own experts. And because the experts' frank exchange will be witnessed by the parties, this diminution in party control may not reduce a party's sense of fair procedure.

Unfortunately, we cannot say much more than this. These suggested changes should reduce adversarialism and remove some impediments to accurate fact-finding, but empirical evidence about these effects is quite limited. What is called for is more experiments involving these procedures: a possibility if some judges would urge the parties to adopt these procedures in their courtroom. [75] Only in these ways will we be able to ascertain whether we can achieve greater accuracy within the overall structure of an adversarial system.

AUTHOR BIOGRAPHIES

NANCY GERTNER is a former Senior United States District Judge of the United States District Court for the District of Massachusetts. Following her retirement from the bench in 2011, she joined the Harvard Law School faculty, most recently serving as Senior Lecturer on Law. She is the author of *In Defense of Women: Memoirs of an Unrepentant Advocate* (2012) and is presently working on her second book, *Incomplete Sentences*.

JOSEPH SANDERS is the A. A. White Professor of Law at the University of Houston Law Center. His recent publications include the volumes *Products Liability: Cases and Materials* (5th ed., with David Fischer, William Powers Jr., Richard Cupp, and Michael Green, 2014) and *Advanced Torts* (3rd ed., with George Christie, 2018).

ENDNOTES

[1] Federal Rules of Evidence, Rule 102.

[2] John W. Thibaut and Laurens L. Walker, *Procedural Justice: A Psychological Analysis* (Hillsdale, N.J.: Erlbaum, 1975); Tom R. Tyler, *Why People Obey the Law* (New Haven, Conn.: Yale University Press, 1990); and E. Allan Lind, Ruth Kanfer, and Christopher P. Earley, "Voice, Control, and Procedural Justice: Instrumental and Noninstrumental Concerns in Fairness Judgments," *Journal of Personality and Social Psychology* 59 (5) (1990): 952 – 959.

[3] As MacCoun notes, a better term perhaps would be procedural fairness, but the term "procedural justice" is so entrenched in the literature that there is no point in adopting a different nomenclature. Robert J. MacCoun, "Voice, Control, and Belonging: The Double-Edged Sword of Procedural Fairness," *Annual Review of Law and Social Science* 1 (2005): 171.

[4] Ibid., 171, 187 – 188; and Ellen S. Cohn, Susan O. White, and Joseph Sanders, "Distributive and Procedural Justice in Seven Nations," *Law and Human Behavior* 24 (5) (2004): 553 – 579.

[5] Other variables play a more important role in other fora, such as interactions with the police. Tom R. Tyler, "What is Procedural Justice? Criteria Used by Citizens to Assess the Fairness of Legal Procedures," *Law and Society Review* 22 (1) (1988): 103, 127 (Table 6).

[6] Joel Brockner and Batia M. Wiesenfeld, "An Integrative Framework for Explaining Reactions to Decisions: Interactive Effects of Outcomes and Procedures," *Psychological Bulletin* 120 (2) (1996): 189, 191; and Robert J. MacCoun, E. Allan Lind, and Tom R. Tyler, "Alternative Dispute Resolution in Trial and Appellate Courts," in *The Handbook of Psychology and Law*, ed. Dorothy K. Kagehiro and William S. Laufer (New York: Springer Verlag, 1992), 95 – 118.

[7] Kees Van den Bos and E. Allan Lind, "Uncertainty Management by Means of Fairness Judgments," *Advances in Experimental Social Psychology* 34 (2002): 1, 25. However, this does not mean that procedures always trump outcomes. When people have a better sense of the outcome of those similarly situated, the procedural fairness associated with greater voice is less important than outcome fairness. A good example of this effect comes from the 9/11 fund for people killed in the World Trade Center. No amount of procedural fairness could overcome the sense of outcome unfairness when family members knew of very different awards to other families based on the standard tort criteria that focus on damages for wrongful death.

[8] Geoffrey Hazard, *Ethics in the Practice of the Law* (New Haven, Conn.: Yale University Press, 1978).

[9] Mirjan R. Damaska, *The Faces of Justice and State Authority: A Comparative Approach to The Legal Process* (New Haven, Conn.: Yale University Press, 1991) [describing the Anglo-American and Continental legal procedures as adversarial and nonadversarial]; Robert A. Kagan, *Adversarial Legalism: The American Way of Law* (Cambridge, Mass.: Harvard University Press, 2001); and Oscar G. Chase, "American 'Exceptionalism' and Comparative Procedure," *American Journal of Comparative Law* 50 (2) (2002): 277.

[10] Max Weber, *The Methodology of the Social Sciences*, ed. and trans. Edward Shils and Henry Fitch (New York: Free Press, 1949).

[11] Full-blown adversarialism is apparently a late-eighteenth-century phenomenon. Stephen Landsman, "The Rise of the Contentious Spirit: Adversary Procedure in Eighteenth Century England," *Cornell Law Review* 75 (3) (1990): 497.

[12] Chase, "American 'Exceptionalism' and Comparative Procedure," 277, 287 [see note 9]. For a valuable discussion of each of these components, see Sam Gross, "Expert Evidence," *Wisconsin Law Review* 1991 (1991): 1113–1232. For articles comparing the use of experts in the United States and other counties, see Gary Edmond and Joëlle Vuille, "Comparing the Use of Forensic Science Evidence in Australia, Switzerland, and the United States: Transcending the Adversarial-Nonadversarial Dichotomy," *Jurimetrics Journal* 54 (3) (2015): 221; Andrew W. Jurs, "Balancing Legal Process With Scientific Expertise: Expert Witness Methodology in Five Nations and Suggestions for Reform of Post-*Daubert* U.S. Reliability Determinations," *Marquette Law Review* 95 (4) (2012): 1329; Andrew J. McClurg, Adem Koyuncu, and Luis Eduardo Sprovieri, *Practical Global Tort Litigation: United States, Germany and Argentina* (Durham: Carolina Academic Press, 2007), 83–97; Petra T. C. van Kampen, "Expert Evidence: The State of the Law in the Netherlands and the United States," in *Adversarial Versus Inquisitorial Justice: Psychological Perspectives on Criminal Justice Systems*, ed. Peter J. van Koppen and Steven D. Penrod (New York: Kluwer, 2003), 210–235; Sven Timmerbeil, "The Role of Expert Witnesses in German and U.S. Civil Litigation," *Annual Survey of International and Comparative Law* 9 (1) (2003): 163, 171; and M. Neil Browne, Carrie L. Williamson, and Linda L. Barkacs, "The Perspectival Nature of Expert Testimony in the United States, England, Korea, and France," *Connecticut Journal of International Law* 18 (2002): 55, 66.

[13] Thibaut and Walker, *Procedural Justice* [see note 2].

[14] Pauline Houlden, Stephen La Tour, Laurens Walker, and John Thibaut, "Preference for Modes of Dispute Resolution as a Function of Process and Decision Control," *Journal of Experimental Social Psychology* 14 (1) (1978): 13–30; and Tyler, "What is Procedural Justice?" 103, 127 (Table 6) [see note 5]. This result has been replicated in European samples. See E. Allan Lind and Tom R. Tyler, *The Social Psychology of Procedural Justice* (New York: Plenum Press, 1988). We are not aware of any studies of this nature that use judges as subjects.

[15] Justin Sevier, "The Truth-Justice Tradeoff: Perceptions of Decisional Accuracy and Procedural Justice in Adversarial and Inquisitorial Legal Systems," *Psychology, Public Policy, and Law* 20 (2) (2014): 212.

[16] Loretta Stalans and E. Allan Lind, "The Meaning of Procedural Fairness: A Comparison of Taxpayers' and Representatives' Views of Their Tax Audits," *Social Justice Research* 10 (3) (1997): 311–331.

[17] John Langbein, "The German Advantage in Civil Procedure," *University of Chicago Law Review* 52 (4) (1985): 823.

[18] See Franklin Strier, "Making Jury Trials More Truthful," *University of California Davis Law Review* 30 (1996): 95; Dan Simon, "Criminal Law at the Crossroads: Turn to Accuracy," *Southern California Law Review* 87 (3) (2014): 421, 436 [adversarial presentation "might indeed be well suited for debates over moral issues, value judgments, and political choices. But zealous argument seems poorly suited as a means for proving objective factual matters."]; Susan Haack, "Epistemology Legalized: Or, Truth, Justice, and the American Way," *American Journal of Jurisprudence* 49 (1) (2004): 43, 49 ["Inquiry is a very different enterprise from advocacy."]; Erin Murphy, "The Mismatch Between Twenty-First-Century Forensic Evidence and Our Antiquated Criminal Justice System," *Southern California Law Review* 87 (3) (2014): 633 ["Ultimately, this Essay argues that almost every aspect of the adversarial process as currently conceived is ill-suited to ensuring the integrity of high-tech evidence."]; and Christopher Slobogin, "Lessons from Inquisitorialism," *Southern California Law Review* 87 (3) (2014): 699 [proposing greater judicial control over the adjudication process and nonadversarial treatment of experts as ways of improving trial accuracy].

[19] Most famous, perhaps, is Wigmore's statement that cross-examination is the "greatest legal engine ever invented for the discovery of truth." John H. Wigmore, *A Treatise on the Anglo-Amer-*

ican System of Evidence in Trials at Common Law*, 3rd ed. (New York: Little, Brown and Company, 1940), 29. More generally, see Monroe H. Freedman, "Judge Frankel's Search For Truth," *University of Pennsylvania Law Review* 123 (1975): 1060; and Lon Fuller, "The Adversary System," in *Talks On American Law*, 2nd ed., ed. Harold J. Berman (New York: Vintage Books, 1971), 34–47. For a useful overview of the arguments defending adversarialism against inquisitorial approaches, see David Alan Sklansky, "Anti-Inquisitorialism," *Harvard Law Review* 122 (2009): 1634.

[20] See Jerome Frank, *Courts on Trial: Myth and Reality in American Justice* (Princeton, N.J.: Princeton University Press, 1949); Thurman W. Arnold, *The Symbols of Government* (New York: Harcourt, Brace and World, 1962); and Marvin E. Frankel, *Partisan Justice* (New York: Hill and Wang, 1978).

[21] Arnold, *The Symbols of Government*, 185 [see note 20].

[22] There is at least one way in which adversarial processes seem to further accuracy. Inquisitorial judges who are in complete control of fact-finding are more likely than judges in an adversarial setting to reach premature conclusions. John Thibaut, Laurens Walker, and E. Allan Lind, "Adversary Presentation and Bias in Legal Decision Making," *Harvard Law Review* 86 (2) (1972): 386. None of the alternatives proposed in this paper would cede complete control over the production of expert testimony to a judge. In any event, when the cases go to trial, a jury, not a judge, will be the ultimate factfinder if a party prefers this.

[23] Anthony Champagne, Daniel Shuman, and Elizabeth Whitaker, "An Empirical Examination of the Use of Expert Witnesses in American Courts," *Jurimetrics Journal* 31 (4) (1991): 375; Daniel W. Shuman, Elizabeth Whitaker, and Anthony Champagne, "An Empirical Examination of the Use of Expert Witnesses in the Courts – Part II: A Three City Study," *Jurimetrics Journal* 34 (2) (1994): 193; and Carol Krafka, Meghan A. Dunn, Molly Treadway Johnson, et al., "Judge and Attorney Experiences, Practices, and Concerns Regarding Expert Testimony in Federal Civil Trials," *Psychology, Public Policy and Law* 8 (3) (2002): 309, 328 (Table 6).

[24] John Thibaut and Laurens Walker, "A Theory of Procedure," *California Law Review* 66 (3) (1978): 541, 541–553; Blair H. Sheppard and Neil Vidmar, "Adversary Pretrial Procedures and Testimonial Evidence: Effects of Lawyer's Role and Machiavellianism," *Journal of Personality and Social Psychology* 39 (2) (1980): 320; and Neil Vidmar and Nancy MacDonald Laird, "Adversary Social Roles: Their Effects on Witnesses' Communications of Evidence and the Assessments of Adjudicators," *Journal of Personality and Social Psychology* 44 (5) (1983): 888.

[25] Max H. Bazerman, George Lowenstein, and Don A. Moore, "Why Good Accountants Do Bad Audits," *Harvard Business Review* 80 (11) (2002): 97, 100–101. See also Lawrence A. Ponemon, "The Objectivity of Accountants' Litigation Support Judgments," *Accounting Review* 7 (3) (1995): 467, 484.

[26] Michael J. Saks and Richard Van Duizend, *The Use of Scientific Evidence in Litigation* (Williamsburg, Va.: National Center for State Courts, 1983); Gross, "Expert Evidence," 1113 [see note 12]; and Jennifer Mnookin, "Expert Evidence, Partisanship, and Epistemic Competence," *Brooklyn Law Review* 73 (587) (2008): 1009.

[27] Emily Pronin, Daniel Lin, and Lee Ross, "The Bias Blind Spot: Perceptions of Bias in Self Versus Others," *Personality and Social Psychology Bulletin* 28 (3) (2002): 369–381.

[28] See, for example, *O'Conner v. Commonwealth Edison Co.*, 807 F. Supp. 1376, 36 Fed. R. Evid. Serv. 589 (C.D. Ill. 1992), judgment aff'd, 13 F.3d 1090, 38 Fed. R. Evid. Serv. 945, 24 Envtl. L. Rep. 20689 (7th Cir. 1994).

[29] Nicholas Scurich, Daniel A. Krauss, Lauren Reiser, et al., "Venire Jurors' Perceptions of Adversarial Allegiance," *Psychology Public Policy and Law* 21 (2) (2015): 161; Sanja Kutnjak Ivkovic and Valerie P. Hans, "Jurors' Evaluations of Expert Testimony: Judging the Messenger and the Message," *Law and Social Inquiry* 28 (2) (2003): 441, 445–446; and Neil Vidmar, "Expert Evidence, the Adversary System, and the Jury," *American Journal of Public Health* 95 (S1) (2005): S137, S137–S142. Complexity also tends to cause jurors to give less weight to expert testimony. Shari Diamond and Jonathan Casper, "Blindfolding the Jury to Verdict Consequences: Damages, Experts and the Civil Jury," *Law and Society Review* 26 (3) (1992): 513, 543 ["lack of clarity,

that is, perceived complexity and difficulty, discourages the jurors from accepting an expert's position, rather than inducing them to accept it"].

30 Jane Goodman, Edith Green, and Elizabeth Loftus, "What Confuses Jurors in Complex Cases," *Trial Magazine*, November 1985, 65, 68. Information about difficulties with eyewitness identifications induced skepticism in experimental jurors but did not improve their ability to distinguish between accurate and inaccurate identifications. See Angela Jones, Amanda Bergold, Marlee Dillon, and Steven Penrod, "Comparing the Effectiveness of *Henderson* Instructions and Expert Testimony: Which Safeguard Improves Jurors' Evaluations of Eyewitness Evidence," *Journal of Experimental Criminology* 13 (1) (2017): 29.

31 See Shari Seidman Diamond, Jonathan Casper, Cami Heiert, and Anna-Maria Marshall, "Juror Reactions to Attorneys at Trial," *Journal of Criminal Law and Criminology* 87 (1) (1996): 17; and Jacqueline Austin, Margaret Bull Kovera, and Michael Lamb, "Cross-Examination Educates Jurors about Missing Control Groups in Scientific Evidence," *Psychology, Public Policy and Law* 21 (3) (2015): 252.

32 Richard E. Petty and John T. Cacioppo, *Communication and Persuasion: Central and Peripheral Routes to Attitude Change* (New York: Springer-Verlag, 1986).

33 Joel Cooper, Elizabeth Bennett, and Holly Sukel, "Complex Scientific Testimony: How Do Jurors Make Decisions?" *Law and Human Behavior* 20 (4) (1996): 379.

34 We should note that everyone both centrally and peripherally processes information. Moreover, as noted below, some peripheral processing is to be desired.

35 Jonathan J. Koehler, N. J. Schweitzer, Michael J. Saks, and Dawn E. McQuiston, "Science, Technology, or the Expert Witness: What Influences Jurors' Judgement About Forensic Science Testimony?" *Psychology, Public Policy and the Law* 22 (4) (2016): 401.

36 Joseph Sanders, "From Science to Evidence: The Testimony on Causation in the Bendectin Cases," *Stanford Law Review* 46 (1) (1993): 1, 47.

37 As one Texas trial attorney pithily remarked, "there have been cases where . . . after I met [the experts], I said 'you know, they are just not going to look too shiny and I need to go get a show dog and do a handoff.'" Joseph Sanders, "The Merits of the Paternalistic Justification for Restrictions on the Admissibility of Expert Evidence," *Seton Hall Law Review* 33 (4) (2003): 881, 913 – 914.

38 Federal Rules of Criminal Procedure, Rule 16 (1) (G).

39 *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

40 David Baltimore, David S. Tatel, and Anne-Marie Mazza, "Bridging the Science-Law Divide," *Dædalus* 147 (4) (Fall 2018). See National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (Washington, D.C.: National Academies Press, 2009), 8; and President's Council of Advisors on Science and Technology, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* (Washington, D.C.: Executive Office of the President, 2016).

41 D. Michael Risinger, "Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?" *Albany Law Review* 64 (1) (2000): 99, 110; and Joseph Sanders, "Applying *Daubert* Inconsistently? Proof of Individual Causation in Toxic Tort and Forensic Cases," *Brooklyn Law Review* 75 (4) (2010): 1367, 1368.

42 See Gross, "Expert Evidence," [see note 12] proposing two ways in which the judiciary could move toward court-appointed experts. Also, here we are thinking of devices such as blue-ribbon juries. Alan Feigenbaum, "Special Juries: Deterring Spurious Medical Malpractice Litigation in State Courts," *Cardozo Law Review* 24 (3) (2003): 1361; and Franklin Strier, "The Educated Jury: A Proposal for Complex Litigation," *DePaul Law Review* 47 (1) (1997): 49.

43 Daniel L. Rubinfeld and Joe S. Cecil, "Scientists as Experts Serving the Court," *Dædalus* 147 (4) (Fall 2018); and Valerie P. Hans and Michael J. Saks, "Improving Judge & Jury Evaluation of Scientific Evidence," *Dædalus* 147 (4) (Fall 2018).

[44] Federal Rules of Evidence, Rule 611.

[45] Murphy, "The Mismatch Between Twenty-First-Century Forensic Evidence and Our Antiquated Criminal Justice System," 633, 665 [see note 18].

[46] Laurens Walker, John Thibaut, and Virginia Andreoli, "Order of Presentation At Trial," *Yale Law Journal* 82 (2) (1972): 216.

[47] Frances P. Kao, Justin L. Heather, Ryan A. Horning, and Martin V. Sinclair Jr., "Into the Hot Tub: A Practical Guide to Alternative Expert Witness Procedures in International Arbitration," *International Law* 44 (3) (2010): 1035, 1038. For additional changes in the ethical rules governing Australian experts, see note 73.

[48] David Sonenshein and Charles Fitzpatrick, "The Problem of Partisan Experts and the Potential for Reform Through Concurrent Evidence," *Review of Litigation* 32 (1) (2003): 1, 59.

[49] Gary Edmond, "Secrets of the 'Hot Tub': Expert Witnesses, Concurrent Evidence, and Judge-Led Law Reform in Australia," *Civil Justice Quarterly* 27 (1) (2008): 51, 56; Gary Edmond, "Merton and the Hot Tub: Scientific Conventions and Expert Evidence in Australian Civil Procedure," *Law and Contemporary Problems* 72 (1) (2009): 159; and Stephen E. Snyder, Daniel Luecke, and John E. Thorson, "Adversarial Collaboration: Court-Mandated Collaboration Between Opposing Scientific Experts in Colorado's Water Courts," *Natural Resources and Environment* 28 (1) (2013): 8.

[50] Edmond, "Secrets of the 'Hot Tub,'" 51, 56 [see note 49].

[51] See Supreme Court of Victoria, "Practice Note No 2 of 2014 Expert Evidence in Criminal Trials," https://www.supremecourt.vic.gov.au/law-and-practice/practice-notes/practice-notes-archive/practice-notes-trial-division-archive-80.

[52] Edmond, "Secrets of the 'Hot Tub,'" 51 [citing as an example the Uniform Civil Procedure Rules, sch. 7 (N.S.W.); see note 49].

[53] Peter Heerey, "Recent Australian Developments," *Civil Justice Quarterly* 23 (2004): 386, 391; and Edmond, "Merton and the Hot Tub," 159, 168–169 [see note 49].

[54] Megan A. Yarnall, "Dueling Scientific Experts: Is Australia's Hot Tub Method a Viable Solution for the American Judiciary," *Oregon Law Review* 88 (1) (2009): 311.

[55] Murphy, "The Mismatch Between Twenty-First-Century Forensic Evidence and Our Antiquated Criminal Justice System," 633 [see note 18].

[56] Sonenshein and Fitzpatrick, "The Problem of Partisan Experts and the Potential for Reform Through Concurrent Evidence," 1, 63 [see note 48].

[57] Erin Murphy, "The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence," *California Law Review* 95 (3) (2007): 721, 753.

[58] Edmond, "Merton and the Hot Tub," 159, 180 [see note 49].

[59] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

[60] National Research Council, *Strengthening Forensic Science in the United States*, 53 [see note 40]. This is so "[f]or a variety of reasons – including the rules governing the admissibility of forensic evidence, the applicable standards governing appellate review of trial court decisions, the limitations of the adversary process, the common lack of scientific expertise among judge and lawyers who must try to comprehend and evaluate forensic evidence – the legal system is ill-equipped to correct the problems of the forensic science community."

[61] President's Council of Advisors on Science and Technology, *Forensic Science in Criminal Courts* [see note 40].

[62] See Risinger, "Navigating Expert Reliability," 99, 149 [see note 41], comparing "the heightened standards of dependability imposed on expertise proffered in civil cases" with "expertise proffered . . . in criminal cases [which] has been largely insulated from any change in pre-*Daubert*

standards or approach." See, for example, Nancy Gertner, "Commentary on the Need for a Research Culture in the Forensic Sciences," *UCLA Law Review* 58 (2011): 89, 790.

63 See, for example, Sara Gordon, "Through the Eyes of Jurors: The Use of Schemas in the Application of 'Plain Language' Jury Instructions," *Hastings Law Journal* 64 (2013): 643, 645 [reforms for improving juror comprehension].

64 Nancy Gertner, "Videoconferencing: Learning Through Screens," *William and Mary Bill of Rights Journal* 12 (3) (2004): 769, 776.

65 Lynne ForsterLee and Irwin Horowitz, "The Effects of Jury-Aid Innovations on Juror Performance in Complex Civil Trials," *Judicature* 86 (4) (2003): 184.

66 B. Michael Dann, Valerie P. Hans, and David H. Kaye, "Can Jury Trial Innovations Improve Juror Understanding of DNA Evidence?" *Judicature* 90 (4) (2007): 152.

67 Lynne ForsterLee, Irwin Horowitz, Elizabeth Athaide-Victor, and Nicole Brown, "The Bottom Line: The Effect of Written Expert Witness Statements on Juror Verdicts and Information Processing," *Law and Human Behavior* 24 (2) (2000): 259.

68 Hans and Saks, "Improving Judge & Jury Evaluation of Scientific Evidence" [see note 43].

69 Nicole L. Mott, "The Current Debate on Juror Questions: 'To Ask or Not to Ask, That Is the Question,'" *Chicago-Kent Law Review* 78 (3) (2003): 1099; and Shari S. Diamond, Mary R. Rose, Beth Murphy, and Sven Smith, "Juror Questions During Trial: A Window into Juror Thinking," *Vanderbilt Law Review* 59 (6) (2006): 1927.

70 See Gross, "Expert Evidence" [see note 12]; Christopher Tarver Robertson, "Blind Expertise," *NYU Law Review* 85 (1) (2010): 174; and Jurilytics, "About Us," https://jurilytics.com/about.

71 Rubinfeld and Cecil, "Scientists as Experts Serving the Court" [see note 43].

72 Joseph Sanders, "Expert Witness Ethics," *Fordham Law Review* 76 (3) (2007): 1539, discusses several of these codes.

73 This code is not contrary to other legal statements about the expert's role. Experts, like all witnesses, take an oath to tell the whole truth. Federal Rule of Evidence 702 envisions that the expert's task is to "help the trier of fact to understand the evidence or to determine a fact in issue." One interpretation of this language is that the law also views the ideal expert witness as a disinterested, unbiased educator. Steven Lubet, a professional responsibility scholar, adopts this position when he argues that "[t]he single most important obligation of an expert witness is to approach every question with independence and objectivity." Steven Lubet, "Expert Witnesses: Ethics and Professionalism," *Georgetown Journal of Legal Ethics* 12 (3) (1999): 465, 467.

74 If a court does choose to adopt a code of ethics, it is important that it be given to the expert at the earliest possible time in the proceedings; ideally at the time the parties list their potential testifying experts. If the court waits until the time of trial to provide the expert with a code of ethics, it may be too late to have much impact on the expert's testimony.

75 Equally important are signals from appellate courts that procedures such as hot-tubbing are not grounds for reversal.

Scientists as Experts Serving the Court

Daniel L. Rubinfeld & Joe S. Cecil

Abstract: Our courts were not designed to consider the increasingly complex scientific and technical evidence needed to resolve contemporary legal disputes. Moreover, when conflicting evidence requires an understanding and interpretation of scientific or technical issues, allowing the parties to control the presentation of evidence places great strain on the judge and jury. This essay describes and evaluates three prototypical procedures that allow courts to appoint scientists and other experts independent of the parties to assist the court: 1) The appointment of an expert to advise the court and the parties regarding a disputed scientific issue by testifying in open court and being cross-examined by the parties; 2) The appointment of a "technical advisor" who assists the judge regarding scientific issues in much the same way that a law clerk assists regarding legal issues; and 3) The appointment of a special master who takes responsibility for the resolution of a portion of the case and prepares a written report for consideration by the court.

DANIEL L. RUBINFELD, a Fellow of the American Academy since 2001, is the Robert L. Bridges Professor of Law and Professor of Economics, Emeritus, at the University of California, Berkeley, and Professor of Law at New York University. He is the author of *Microeconomics* (9th ed., with Robert S. Pindyck, 2017) and *Economics Models and Economics Forecasts* (with Robert S. Pindyck, 1998).

JOE S. CECIL is Senior Research Associate at the Federal Judicial Center. He has published several reports and articles on court-appointed experts, including in journals such as *American Journal of Public Health*.

Our courts were not designed to consider the increasingly complex scientific and technical evidence needed to resolve contemporary legal disputes. The common law tradition of the United States relies on the litigating parties to structure the presentation of evidence by selecting witnesses and allows them in some measure to shape their evidence presentation to their own advantage. While there are limits on the extent to which this can be done with ordinary witnesses, there is far greater leeway in shaping the evidence presented by expert witnesses. Indeed, if a party does not like what one retained expert has to say, the party need not call that expert and can instead present another expert whose testimony better supports the party's case. In most instances, the opposing side and the factfinder will not even know that another expert had been consulted. Similarly, cross-examination by the opposing party is supposed to identify weaknesses in opposing witness testimony by revealing inconsistencies, showing flaws in opportunities to observe, and revealing biases and other motives to deceive. Juries and judges are expected to understand

doi:10.1162/DAED_a_00526

the points being made. Experts, however, typically know considerably more than opposing counsel about the matters they discuss in their testimony, and jurors and judges typically know even less. Some experts are experienced witnesses selected in part for their proven success in withstanding cross-examination and persuading judges and jurors to their side. Others may be novices, inexperienced in giving testimony, who communicate poorly to the jury or are easily flustered by a cross-examiner, even if their science is sound.

After hearing the evidence, judges and juries render judgments, sorting out, to the best of their abilities, the conflicting evidence presented by the parties. This common law tradition ensures that the parties are given a full and fair opportunity to present their strongest case and appears to work well when conflicting evidence involves issues within the common knowledge and experience of the judge and jury. But when conflicting evidence requires an understanding and interpretation of scientific or technical issues, allowing the parties to control the presentation of evidence places great strain on both the judge and jury. Few judges and juries arrive in the courtroom with the knowledge and experience necessary to resolve patent disputes over new genetic technologies or antitrust disputes involving the proper specification of a commercial market. Hence the parties' expert witnesses are critical to understanding and resolving the scientific and technical issues that may lie at the heart of the dispute; weaknesses in the adversary system's capacity to deal with experts threatens both accuracy and justice.

Judges have an affirmative duty to ensure that expert testimony is scientifically valid and reliable, and modern courts often face motions to exclude proffered expert testimony because it lacks a proper foundation in scientific practice.[1] Yet without some form of assistance, judges and juries are unlikely to know if such testimony is consistent with the scientific consensus, and they may have great difficulty determining the proper weight to give the views of the opposing experts. If the parties' experts are unable or unwilling to educate the court regarding their areas of agreement and disagreement, the resulting court decision may be at odds with the current understanding of the scientific community and, in the extreme case, may be based on methods or theories for which there is no respectable scientific support.

Judges and juries are not helpless when faced with complex scientific evidence, however. As indicated by Nancy Gertner and Joseph Sanders in their essay in this volume, over the years judges have developed procedural techniques that have strengthened their ability to assess the foundation of expert testimony and clarify its complexities.[2] Nevertheless, in some cases, evidence is so complex that commonly available procedural devices are inadequate to provide judges and juries with a sufficient understanding of the conflict to allow a reasoned and principled decision. It is in those extraordinary circumstances that a judge should consider going beyond the common law tradition and seek the assistance of an expert appointed by the court and not sponsored by the parties.[3]

Judges and attorneys are well aware of the problems that expert testimony presents in a common law system.[4] A 1999 survey of federal judges and attorneys found that, in their view, the most frequent problem by far with expert testimony is that "[e]xperts abandon objectivity and become advocates for the side that hired them."[5] The third and fourth most frequent problems are that "[e]xpert testimony appears to be of questionable validity or reliability" and "[c]onflict among experts [is presented] that defies reasoned assessment."[6] Expert witnesses' abandonment of objectivity, becoming advocates for a party, and offering

invalid scientific testimony are facilitated, if not encouraged, by the parties' ability to select and shape the expert testimony that most favors their interests. While there are benefits to giving parties their choice of experts and a role in presenting their experts' testimony, it is difficult to defend practices that so frequently limit the opportunity for judges and juries to reach a reasoned and principled resolution of a dispute among scientific experts.

The most obvious alternative to the common law practice is to submit disputes regarding the scientific evidence to one or more independent scientists for resolution. In fact, in 1967, physicist Arthur Kantrowitz suggested the development of a "science court" composed of scientists and other experts that would resolve scientific conflicts that arose in the context of public discourse, including in litigation.[7] The science court proposal sought to distinguish scientific issues from other disputed issues in the litigation and to submit the scientific issues to an advisory committee that would resolve the scientific conflict. These findings would then be incorporated into the legal proceeding and guide resolution of the remaining nonscientific issues.

The science court proposal was widely discussed in the 1970s, and was the focus of a task force report of the White House Office of Science and Technology Policy.[8] Some critics questioned whether it was possible to separate scientific facts from the legal, political, and moral issues that may arise in disputes. Others thought such a procedure might stifle scientific debate. Despite the endorsement of the task force report and the support of a number of scientific organizations, the proposed science court was never tested.[9] Changes in political administrations and lack of funding doomed the proposal, and attention turned to other forms of adjudicating disputes over scientific issues.

In the succeeding years, other procedures to allow independent scientists to assist the courts have been developed. In this essay, we describe three prototypical procedures that allow courts to appoint scientists and other experts who are independent of the parties to assist the court. The first and most widely known procedure is described in Rule 706 of the Federal Rule of Evidence and related state court rules. This rule allows the court to appoint an expert to advise the court and the parties on scientific matters by testifying in open court subject to cross-examination by the parties, in the same way party experts can be examined. The second procedure is court appointment of a "technical advisor" who assists the judge regarding scientific issues in much the same way that a law clerk assists regarding legal issues. This alternative is especially useful for cases in which a judge must master a complex body of knowledge in order to render a decision. Last, a court may appoint a special master who takes initial responsibility for the resolution of the portion of the case involving scientific issues and prepares a written report for consideration by the court. Special masters are typically appointed to deal with accounting tasks or the computation of damages in complex cases, but in extraordinary circumstances, scientists or engineers have served as special masters to assist the court in resolving disputes over patent claims or to supervise discovery of technical information, such as computer source code. Each of these three alternatives is a compromise intended to strengthen the ability of the judge and jury to resolve disputes over scientific issues within the broad confines of our common law tradition.[10]

Experts appointed in accordance with Federal Rule of Evidence 706 are chosen by the judge following consultation with both parties, with expert fees and other costs typically borne equally by both sides.

The principal role of court-appointed experts is to give expert testimony.[11] The expert is expected to prepare an expert report (oral or written) based on the available evidence and, as the case proceeds to trial, to be available for depositions and cross-examination by counsel for both parties. Appointment of experts by the federal courts under Rule 706 is rare. A 1993 study by the Federal Judicial Center found that only half of the federal district court judges had ever appointed a Rule 706 expert, and only half of those had appointed an expert on more than one occasion.[12] Nevertheless, an appropriately managed expert role can be highly beneficial to the court, although there remain some concerns.[13]

On the positive side, court-appointed experts can clarify fundamental issues addressed by the parties' experts and diminish the extent to which the court must rely on their testimony, perhaps securing agreement on undisputed facts from these experts or otherwise eliminating the need for them. Narrowing the role of the partisan experts may not only save costs; it may also diminish any concerns that the judge or jury must rely on "hired guns" who are paid to take a position. More important, the conclusions of a court-appointed expert may encourage settlement or provide better grounding for the outcomes in cases that go to trial.

Court-appointed experts facilitate settlements by offering a more "neutral" evaluation of the issues in the case, decreasing the possibility that the parties will cling to extreme positions. Whether the expert suggests an appropriate outcome or simply presents an accurate characterization of the views of the two parties, the expected value of a case is likely to be different and more certain than it would be without the expert. These changes will typically enhance the likelihood that a case will settle.[14]

Among the roles the court may define for an appointed expert are any or all of the following. First, the expert can brief the trier of fact (the judge or the jury) on foundational scientific and technical concepts that underlie the dispute between the parties. Second, the expert can assist the judge in determining whether the party expert's testimony is sufficiently grounded in sound science to be admissible by impartially summarizing the available research and informing the court of the extent of scientific consensus on an issue of import. Third, the expert can provide a conceptual framework that aids the judge and jury in assessing the validity of the different opinions offered by the parties' experts. Finally, and most significant, the presence of a court-appointed expert can change the incentives of the partisan experts for the parties. Knowing that their work will be scrutinized by a highly qualified neutral expert appointed by the court, partisan experts are likely to give more focused testimony that is more firmly based on a solid scientific foundation.[15] This will not only increase the likelihood of settlement, but, when cases do not settle, can also lead to scientifically sounder decisions by the trier of fact.

There is, of course, no free lunch with the appointment of a neutral expert. First and most obvious is the increased costs in time and effort to the parties and to the court itself. Depending on the area of needed expertise, there may be substantial controversy as to the set of appropriate skills and qualifications of potential appointees. The court must then identify a suitable expert, screen the expert for conflicts of interest, and secure compensation for the expert from the parties. Additionally, it takes time to instruct the court-appointed expert on the tasks the expert must accomplish and to allow the appointed expert to become familiar with the issues in dispute. The parties must also spend time considering and responding to the findings and testimony of the appointed expert. These additional time and effort costs, along with

the expert's fees, translate into dollar costs that must be paid by the parties. Depending on what is in dispute, it may be difficult for a judge to justify these costs in a given case, and even harder for lawyers to justify the presence of court-appointed experts to their clients, especially when participation by an appointed expert could weaken their case.

Second, court-appointed experts may undermine the authority of the judge or jury by acquiring an unentitled "aura of infallibility." In the extreme, if not appropriately cabined, the expert could take over the judicial role by framing the resolution of scientific and technical issues in ways that intrude on the authority of the judge to interpret the law. While this is an unlikely possibility, the prospect that the lawyers for the parties may lose some control over presentation of their case may be a valid fear. Judges and juries are likely to discount the views of the parties' experts when their views are in conflict with those of an expert appointed by the court, since the views of the court-appointed expert are likely to be regarded as free of partisan bias. But partisan bias is not the only bias that may affect an expert's testimony. In some areas, disciplines are divided about the weight or quality of available evidence or the import of theory. A court-appointed expert may belong to one school of thought rather than another and this debatable intellectual bias may shape what the jury hears.

If, however, the court-appointed expert's role is well-defined and appropriately cabined, we do not see these concerns as outweighing the opportunity for factfinders to reach more informed judgments based on sound scientific concepts.[16] For example, in *Monolithic Power Systems v. O2 Micro International*, the United States Court of Appeals for the Federal Circuit ruled that the district court did not abuse its discretion in appointing an expert under Federal Rule 706.[17] With respect to the appointment of electrical engineer Enrico Santi as a neutral expert, the court noted:

> The predicaments inherent in court appointment of an independent expert and revelations to the jury about the expert's neutral status trouble this court to some extent. Courts and commentators alike have remarked that Rule 706 should be invoked only in rare and compelling circumstances. *In re Joint E. & S. Dists. Asbestos Litig.*, 830 F. Supp. 686, 693 (E.D.N.Y. 1993) (noting that "use of Rule 706 should be reserved for exceptional cases in which the ordinary adversary process does not suffice"); Wright, *supra*, § 6302 ("Rule 706 powers are properly invoked where the issues are complex and the parties' experts have presented conflicting testimony that is difficult to reconcile or have otherwise failed to provide a sufficient basis for deciding the issues."). However, under Ninth Circuit law, district courts enjoy wide latitude to make these appointments. This court perceives no abuse of discretion in this case where the district court was confronted by what it viewed as an unusually complex case and what appeared to be starkly conflicting expert testimony.[18]

As an illustration of the use of a court-appointed expert, consider the Glass Containers Antitrust Litigation.[19] The plaintiffs brought an antitrust price-fixing case against the major manufacturers of glass containers and sought to certify a single, national class that included all direct purchasers of manufactured glass containers. In order to certify such a class, the plaintiff was required to show that the class members suffered similar damages.[20] In opposition to class certification, the defendants' economic expert argued that there was substantial price variation among different types of glass containers and that the price variation was evidence that any harm that might have been suffered by putative class members would have varied substantially among individuals. After fur-

ther discovery into manufacturing costs issues that both sides' experts had avoided, the court-appointed expert and coauthor of this essay Daniel Rubinfeld found that variations in glass container pricing were best explained by a series of pricing equations, each of which explained pricing for a different type of container (such as a wine bottle or a pickle jar). Ultimately, District Court Judge Ilana Rovner followed Rubinfeld's analysis by certifying a set of subclasses made up of a variety of types of glass containers. Although the litigation continued for some time, the eventual settlement of the case was driven in part by Rubinfeld's expert report.

Several aspects of the role of the court-appointed expert in this case are worth noting. First, the plaintiff's filing of a writ of mandamus, which questioned Judge Rovner's plan to appoint a court-appointed expert, was rejected by the Seventh Circuit Court of Appeals.[21] Second, Judge Rovner chose to have no *ex parte* contact with the court-appointed expert, a choice that avoided potential claims of bias on the part of the judge. Third, there is little doubt that, had the appointment of the neutral expert been made earlier in the litigation, the reports of the parties' experts would have been more focused on the central issues in the case.[22] Fourth, because the case involved substantial potential damages, the case was able to support focused expert discovery by the court-appointed expert, followed by his written report and oral testimony. Both sides responded at length in written replies, but chose to only minimally cross-examine the neutral expert's oral testimony at a court-directed hearing. This may have reflected the parties' fear of irritating Judge Rovner, who had made the appointment in the first place.

Another example is the extensive use Judge Kimba Wood made, over the plaintiff's objection, of court-appointed expert economist Alfred Kahn in *State of New York v. Kraft General Foods, Inc.*[23] At issue was the legality of Kraft's planned acquisition of Nabisco's cereal assets. At Judge Wood's request, the expert economists for both parties and Kahn appeared in court throughout the thirty-day bench trial. Not only opposing counsel but also Judge Wood and Kahn cross-examined each of the parties' experts on each of the fundamental merger issues in the case: market definition, market power, and competitive effects. In the end, Kahn testified as to his views of each of these issues and, in turn, was available for cross-examination by both parties.

Kahn's role (and the role of his associates) in this matter was no doubt influential. While it was costly to the parties, it had the effect of focusing the testimony of the experts on critical issues and clarifying portions of the experts' testimonies for Judge Wood. Perhaps due in part to the role of the court-appointed expert, the State of New York chose not to appeal Judge Wood's opinion in favor of Kraft.[24]

A more amorphous role for the neutral expert is as advisor to the court on technical issues. Although not authorized by a specific rule, judges have relied on the court's inherent authority to appoint technical advisors, who function more like law clerks than like testifying expert witnesses.[25] Technical advisors may function as "a sounding board" that can help the jurist to educate him- or herself in the jargon and theory that the parties' experts or presentations have referenced and to think through critical technical problems.[26] There are a number of advantages associated with this more limited role. First, the job is likely to be less time-consuming and expensive than the role of a testifying expert. Indeed, a technical advisor need not submit to a deposition or cross-examination by counsel for the parties.

Second, the expert can provide useful information for a court conducting a *Daubert*

hearing where the methodology or qualifications of one or more of the party-selected experts is challenged. In this setting, the technical advisor is aiding the judge in his or her role as gatekeeper. This may result in judicial rulings that either promote settlement or, on occasion, leave one party (almost always the plaintiff) without a case, saving the court as well as the parties time and money (even if one party is deeply disappointed with the result). Of course, there remains the danger that the expert will be unduly influential in determining the outcome of the case, as when a court precludes expert testimony and dismisses the case in a situation where a different neutral expert might have suggested to the judge that the *Daubert* hurdle had been cleared.

Third, the advisor can inform the judge about issues relating to data and methods, as well as help the judge devise plans for jointly agreed upon data sets or methodological approaches. Furthermore, because no testimony is planned, the court has substantial flexibility in its use of the expert. Indeed, if there is no testimony, the Rules of Evidence will not come into play. We see here a potentially valuable instructional role for the expert, for example, giving an unbiased tutorial to the court.

As with Rule 706 experts, there are concerns even when experts fill only this limited role. First, there is a possibility that the expert may go beyond the judge's remit and inappropriately influence the judge. Second, the expert's opinion may do little to help the parties converge on a settlement range and, in the end, not be cost-effective. Third, a substantial amount of information may be conveyed to the judge without the knowledge of or scrutiny by the parties. Indeed, there is no cross-examination of the technical advisor, contrary to the ideals of the adversarial system. For these reasons, a leading medical malpractice case warned that technical advisors should be "hen's-teeth rare" and should be used "only where the trial court is faced with problems of unusual difficulty, sophistication, and complexity, involving something well beyond the regular questions of fact and law with which the judges must routinely grapple."[27]

The difficult role of the technical advisor is demonstrated in the contentious case of *Association of Mexican-American Educators v. State of California*.[28] A number of minority educators filed a class-action lawsuit against the state, challenging the defendants' use of the California Basic Educational Skills Test (CBEST) as a requirement for certification to teach in California public schools. As part of this challenge, the plaintiffs questioned the validity of the test. After years of discovery concerning the development and application of the test, the court found that it needed guidance regarding a number of complex technical issues. The court then appointed Stephen P. Klein, an expert in test validation, as a technical advisor to the court. In the order appointing Klein, the court noted that it had come "face to face with many prickly problems requiring expertise in the esoteric fields of education and psychometrics including knowledge of theories about educational measurement and testing, cognitive psychology, statistics, and other fields pertaining to the CBEST and other cases." Klein was asked to review all of the expert testimony submitted in the case and to "confer *ex parte* with the Court from time to time." After considering the evidence, assisted by the assessments of Klein, the court found no violation of the Civil Rights Act. The plaintiffs objected to the *ex parte* nature of the court's communication with Klein. On appeal, the Ninth Circuit Court of Appeals upheld the appointment of the technical advisor and, in the absence of some indication of impropriety, did not object to the *ex parte* nature of the communication.[29]

The Federal Rules allow broad grants of authority to special masters to aid the

court. Under Federal Rule of Civil Procedure 53, the special master may perform any duties "consented to by the parties." Absent such consent, the appointed master may still address pretrial and posttrial matters that will assist the judge. In cases decided without a jury, the special master may go further and hold trial proceedings and make or recommend findings of fact regarding damages, the results of an accounting, or when there is an exceptional condition that warrants such assistance. In making such appointments, the court is expected to bear in mind the likely cost that will be imposed on the parties and any possibility of unreasonable delay.[30]

There are a number of benefits to the court and the parties flowing from a Rule 53 appointment. First, the appointment offers more flexibility for a judge than the appointment of a testimonial expert under Rule 706. The master's role might be sufficiently broad so as to provide judges with help for time-consuming tasks and, as a result, can shorten the litigation. Second, compared to the judge, a master who is responsible for only one case can more effectively deal with specific tasks such as monitoring long-term compliance or consent decrees. Third, a master can mitigate the possibility of judicial biases. Judges must pass on the admissibility of evidence, and even though they rule evidence inadmissible, their knowledge of the inadmissible evidence may color their decisions. A master can winnow admissible from inadmissible evidence so that judges are not exposed to the former. Fourth, expert masters, unlike most judges, have the time and special knowledge needed to receive and sort through reams of evidence to identify and organize the information that is of greatest consequence to a just outcome of the litigation.

By performing a variety of tasks, including damages calculations, the master can assist the judge with complex issues, while freeing up judicial resources, an important goal given the complexities of e-discovery in complex litigation. Moreover, with more time to focus on the qualifications of experts and determine where scientific consensus lies, a master can enhance the quality of *Daubert* evaluations, which can lead to quicker settlements or the dismissal of a case.

As with the approaches discussed above, there are, however, potential downsides to the use of masters. One is the cost to the parties who generally must pay the special master's fees. In addition, adding a role for a master into litigation can create delays that would not exist if the case were handled throughout by the trial judge, particularly since masters can take evidence sporadically, while, once a judicial trial begins, proceedings are most often more or less continuous. Moreover, the need to pay masters and their capacity to probe more deeply into a matter than a judge might on issues that arose pretrial or that might occur in party-controlled litigation can raise parties' costs, perhaps substantially. Finally, judicial authority and suggestions of judicial leanings are often an important incentive to settlement. Similar behavior by masters may have less clout.

Medtronic Sofamor Danek, Inc. v. Michelson offers an example of how the appointment of a scientist as special master may aid the court in dealing with discovery requests for complicated technical data.[31] The plaintiff sought discovery of electronic data that related to trade secrets, patents, and trade information in the field of spinal fusion medical technology. The requested information was contained on 996 network backup tapes, which included, among other things, the plaintiff's electronic mail and an estimated three hundred gigabytes of other, nonbacked-up electronic data. In light of the enormous amount of data that was to be procured, the judge appointed an expert trained in computer science and relat-

ed technologies as special master to oversee the electronic records production and to review the data files produced in response to discovery requests. The special master's duties included "making decisions with regard to search terms; overseeing the design of searches and the scheduling of searches and production; coordinating deliveries between the parties and their vendors; and advising both parties, at either's request, on cost estimates and technical issues."[32]

This case also offers an example of the need for care in defining the duties of and in monitoring the special master to avoid an improper delegation of judicial authority to a nonjudicial official. At the conclusion of five months of discovery, the plaintiff challenged the decision of the special master to withhold some of the tapes after finding that they did not contain deleted files. After the special master denied the plaintiff's request, the plaintiff sought review by the court. The court determined that the special master had exceeded his authority in denying the request, since the order of appointment did not include the duty of "making determinations as to whether Medtronic could be compelled to produce deleted files and e-mails."[33]

While there remain concerns about the use of court-appointed experts, whatever their capacity, we find the case for a more expansive role to be compelling. The more frequent use of such "neutral" experts seems particularly desirable given the increasing complexity of litigation in areas ranging from antitrust and intellectual property law to employment discrimination. We find particularly noteworthy the support this position has received from Supreme Court Justice Stephen Breyer and, more recently, from Circuit Court Judge Richard Posner.[34] According to Judge Posner,

> Turning to the technical statistical evidence …we recommend that the district judge use the power that Rule 706 of the Federal Rules of Evidence expressly confers upon him to appoint his own expert witness, rather than leave himself and the jury completely at the mercy of the parties' warring experts. The main objection to this procedure and the main reason for its infrequency are that the judge cannot be confident that the expert whom he has picked is a genuine neutral. The objection can be obviated by directing the party-designated experts to agree upon a neutral expert whom the judge will then appoint as the court's expert. The neutral expert will testify (as can, of course, the party-designated experts) and the judge and jury can repose a degree of confidence in his testimony that it could not repose in that of a party's witness. The judge and jurors may not understand the neutral expert perfectly but at least they will know that he has no axe to grind, and so, to a degree anyway, they will be able to take his testimony on faith.[35]

In closing, we reiterate that the obligation of the court remains to ensure that the expert, court-appointed or otherwise, provides the factfinder with maximally understandable evidence.

AUTHORS' NOTE

The authors thank Judge Jack Weinstein, Shari Diamond, and Rick Lempert for their helpful comments. The views expressed herein are those of the authors and do not necessarily represent the views of the Federal Judicial Center.

ENDNOTES

[1] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

[2] Nancy Gertner and Joseph Sanders, "Alternatives to Traditional Adversary Methods of Presenting Scientific Expertise in the Legal System," *Dædalus* 147 (4) (Fall 2018).

[3] The notion of a scientist appointed to aid the court may be a departure from the common law tradition, but experts appointed by the courts are a common practice in Europe and other court systems that rely on a civil law tradition. See James G. Apple and Robert P. Deyling, *A Primer on the Civil Law System* (Washington, D.C.: Federal Judicial Center, 1995). In civil law courts, the judge plays a more active role in developing the evidence in a case and may direct the parties to provide the information deemed necessary to resolve the case. Under this civil law "inquisitorial" system, the court may appoint expert witnesses on its own initiative, and some courts maintain a list of such experts who can be called upon as needed to assist the court.

[4] Concern over the ability of courts to deal with complex scientific and technical evidence is long-standing. See, for example, Learned Hand, "Historical and Practical Considerations Regarding Expert Testimony," *Harvard Law Review* 15 (1) (1901): 40 – 58 [discussing the "anomaly" of asking a lay jury to resolve a dispute between expert witnesses], cited in Harold L. Korn, "Law, Fact & Science in the Courts," *Columbia Law Review* 66 (6) (1966): 1080 – 1116. Korn noted: "The unprecedented challenge of keeping the legal system attuned to the current pace of scientific advance suggests that [expert testimony] may now require a more broadly based kind of scrutiny than it has traditionally received." Ibid.

[5] Carol L. Krafka, Megan A. Dunn, Molly Treadway Johnson, et al., "Judge and Attorney Experiences, Practices, and Concerns Regarding Expert Testimony in Federal Civil Trials," *Psychology, Public Policy, and Law* 8 (3) (2002): 328.

[6] Ibid. The second most frequently cited problem is "[e]xcessive expense of party-hired experts."

[7] Arthur Kantrowitz, "Proposal for an Institution for Scientific Judgment," *Science* 156 (3776) (1967): 763 – 764. The history of various proposals for a science court are summarized in Andrew W. Jurs, "Science Court: Past Proposals, Current Considerations, and a Suggested Structure," *Virginia Journal of Law and Technology* 15 (2010): 1 – 42. See also Arthur Kantrowitz, "Controlling Technology Democratically," *American Scientist* 63 (5) (1975): 505 – 509; and Justin Sevier, "Redesigning the Science Court," *Maryland Law Review* 73 (3) (2014): 770 – 835.

[8] Task Force of the Presidential Advisory Group on Anticipated Advances in Science and Technology, "The Science Court Experiment: An Interim Report," *Science* 193 (1976): 653.

[9] For a discussion of an attempt to use the "science court" to resolve a dispute over placement of a high-voltage power line, see Jurs, "Science Court" [see note 7].

[10] When the judge is the trier of fact, he or she might also find it useful to have the experts serve an educational function by offering relevant substantive viewpoints prior to or at the beginning of the formal judicial proceedings. These proceedings could, in principle, include a role for a court-appointed expert.

[11] According to Rule 706(a): "The court may appoint any expert witnesses agreed upon by the parties, and may appoint expert witnesses of its own selection. . . . A witness so appointed shall be informed of the witness' duties by the court in writing, a copy of which shall be filed with the clerk, or at a conference in which the parties shall have opportunity to participate. A witness so appointed shall advise the parties of the witness' findings, if any; the witness' deposition may be taken by any party; and the witness may be called to testify by the court or any party. The witness shall be subject to cross-examination by each party, including a party calling the witness." Federal Rules of Evidence, Rule 706(a). For a more detailed discussion of Rule 706, see Jack B. Weinstein, Norman Abrams, Scott Brewer, and Daniel Medwed, *Evidence: Cases and Materials* (Eagan, Minn.: Foundation Press, 2017), chap. 7.

[12] Joe S. Cecil and Thomas E. Willging, *Court-Appointed Experts: Defining the Role of Experts Appointed under Federal Rule of Evidence 706* (Washington, D.C.: Federal Judicial Center, 1993). One no-

table instance of such an appointment involved the appointment of a multidisciplinary panel of experts to aid the court in resolving competing claims of harm caused by leakage from silicone gel breast implants. See Laural L. Hooper, Joe S. Cecil, and Thomas E. Willging, "Assessing Causation in Breast Implant Litigation: The Role of Science Panels," *Law and Contemporary Problems* 64 (4) (2001): 139 – 189. Approximately twenty-seven thousand cases alleging injuries from defective implants were consolidated in one federal jurisdiction. When it became clear that the key underlying dispute involved conflicting expert testimony regarding whether the leaking breast implants caused systemic connective tissue diseases such as lupus, the judge appointed a multidisciplinary panel of biomedical experts to review the scientific literature and prepare a written report that could be used in subsequent trials nationwide. While the expert panel report indicating no causal relationship appeared to be successful in moving the parties toward a resolution of this issue, the procedure also proved to be more costly than expected and the federal courts have not agreed to pay for another such expert panel. For the views of the scientists who served on the panel, see Barbara S. Hulka, Nancy L. Kerkvliet, and Peter Tugwell, "Experience of a Scientific Panel Formed to Advise the Federal Judiciary on Silicone Breast Implants," *New England Journal of Medicine* 342 (11) (2000): 812 – 815.

[13] Court-appointed experts in state court cases are quite rare. Judge John Wiley points out that while experts can be appointed by California judges under California Evidence Code Section 730, that option has rarely been used. See John Shepard Wiley, "Taming Patent: Six Steps for Surviving Scary Patent Cases," *UCLA Law Review* 50 (2003): 1413 – 1482.

[14] See, for example, J. J. Prescott and Kathryn Spier, "A Comprehensive Theory of Settlement," *NYU Law Review* 91 (1) (2016): 60 – 143.

[15] The Rule 706 advisory committee notes that "The ever-present possibility that the judge may appoint an expert in a given case must inevitably exert a sobering effect on the expert witness."

[16] We do note, however, that a small number of states have removed Rule 706 from their evidence code. See Edward J. Imwinkelried, "The Court Appointment of Expert Witnesses in the United States: A Failed Experiment," *International Journal of Medicine and Law* 8 (1989): 601 – 609. For a survey of state court practices, see Stephanie Domitrovich, Mara L. Merlino, and James T. Richardson, "State Trial Judge Use of Court Appointed Experts: Survey Results and Comparisons," *Jurimetrics Journal* 50 (3) (2010): 371 – 389.

[17] *Monolithic Power Systems v. O2 Micro International, Ltd.*, 726 F.3d 1359 (2013).

[18] *Monolithic Power Systems, Inc., et al. v. O2 Micro International, Ltd.*, 558 F.3d 1341 (Fed. Cir. 2009). "*See, e.g., Walker*, 180 F.3d at 1071 (finding no abuse of discretion in Rule 706 appointment where the scientific evidence was 'confusing and conflicting' and the appointment 'assist[ed] the court in evaluating contradictory evidence about an elusive disease of unknown cause')." Ibid.

[19] *Superior Beverage Company v. Owens-Illinois, Inc.*, No. 83 C 512, 1989 U.S Dist. LEXI 6662 (N.D. Ill. June 5, 1989). The role of coauthor Daniel Rubinfeld as the expert is described in Bret Dickey and Daniel Rubinfeld, "Antitrust Class Certification: Towards an Economic Framework," *NYU Law Annual Survey of American Law* 66 (3) (2011): 459 – 486.

[20] Federal Rule of Civil Procedure 23(a)(2). See also *Comcast Corp., et al. v. Behrend, et al.*, 133 S.Ct. 1426 (2013).

[21] Judge Rovner was later appointed to the Circuit Court.

[22] This is based on two informal conversations, one between Daniel Rubinfeld and the plaintiff's expert, business administration scholar Albert Madansky, and one between Rubinfeld and the defendant's expert, economist Franklin Fisher.

[23] *State v. Kraft General Foods, Inc.*, 926 F. Supp. 321 (S.D.N.Y. 1995).

[24] While courts have historically appointed medical experts to give technical advice, the more recent growth in the use of court-appointed experts has been in the intellectual property area. A recent example is Judge Richard Posner's appointment of antitrust expert Gregory Sidak as a neutral expert in the evaluation of intellectual property damages. For a detailed analysis of his role of court-appointed expert in *Brandeis University v. East Side Ovens, Inc.*, Nos. 1:12-cv-01508,

1:12-cv-01509, 1:12-cv-01510, 1:12-cv-01511, 1:12-cv-01512, 1:12-cv-01513 (N.D. Ill. 2013), see J. Gregory Sidak, "Court-Appointed Neutral Economic Experts," *Journal of Competition Law and Economics* 9 (2) (2013): 359 – 394.

[25] *Ex Parte Peterson*, 253 U.S. 300, 40 S.Ct. 543, 64 L.Ed. 919 (1920). "Courts have (at least in the absence of legislation to the contrary) inherent power to provide themselves with appropriate instruments required for the performance of their duties. This power includes authority to appoint persons unconnected with the court to aid judges in the performance of specific judicial duties, as they may arise in the progress of a cause."

[26] *Reilly v. United States*, 863 F.2d 149, 158 (1st Cir. 1988). See also "Improving Judicial Gatekeeping: Technical Advisors and Scientific Evidence," *Harvard Law Review* 110 (4) (1997).

[27] *Reilly v. United States*, 157 [see note 26].

[28] *Association of Mexican-American Educators v. State of California*, 231 F.3d 572 (2000).

[29] In dissent, Judge Tashima suggested a number of safeguards for ensuring that technical advisors provided assistance in a fair and neutral manner. Ibid., 611 – 614. More recently, Magistrate Judge James Orenstein (Eastern District of New York) appointed economics law scholar Alan Sykes to advise the court with respect to economic issues in connection with a proposed settlement in a large complex credit card case brought by Target and other retailers against Visa, Mastercard, and Discover. The retailers had some concerns relating to discovery, but the parties did not object to the appointment and the proposed settlement was approved by the court. The case is *In re Payment Card Interchange Fee and Merchant Discount Antitrust Litigation*, Case No. 1:05-md-01720 (E.D.N.Y. January 10, 2014).

[30] For an overview of the historical use of Rule 53, see Thomas E. Willging, Laural L. Hooper, Marie Leary, et al., *Special Masters' Incidence and Activity: Report to the Judicial Conference's Advisory Committee on Civil Rules and Its Subcommittee on Special Masters* (Washington, D.C.: Federal Judicial Center, 2000).

[31] *Medtronic Sofamor Danek, Inc., v. Michelson*, 229 F.R.D. 550 (2003). See also Shira A. Scheindlin and Jonathan M. Redgrave, "Special Masters and E-Discovery: The Intersection of Two Recent Revisions to the Federal Rules of Civil Procedure," *Cardozo Law Review* 30 (2) (2008): 372 – 374.

[32] Ibid., 559.

[33] *Medtronic Sofamor Danek, Inc. v. Michelson*, Nos. 01-2373, 03-2055, 2004 WL 2905399 (W.D. Tenn. May 3, 2004).

[34] Justice Breyer's view was promulgated in *General Electric Company v. Joiner*, 522 U.S. 136, 148 (1997). According to Justice Breyer, "[A] judge could better fulfill this gatekeeper function if he or she had help from scientists. Judges should be strongly encouraged to make greater use of their inherent authority . . . to appoint experts. . . . Reputable experts could be recommended to courts by established scientific organizations, such as the National Academy of Sciences or the American Association for the Advancement of Science."

[35] *In re High Fructose Corn Syrup Antitrust Litig.*, 295 F.3d 651, 665 (7th Cir. 2002). This reference and the Breyer reference are described in an insightful article in support of the use of court-appointed experts; see Judge Bradford H. Charles, "Rule 706: An Underutilized Tool to be Used when Partisan Experts Become 'Hired Guns,'" *Villanova Law Review* 60 (2015): 941 – 954. For evidence of additional judicial support, see Shira Scheindlin, "We Need Help: The Increasing Use of Special Masters in Federal Court," *DePaul Law Review* 58 (2009): 479 – 486.

Improving Judge & Jury Evaluation of Scientific Evidence

Valerie P. Hans & Michael J. Saks

Abstract: The role of the expert witness in trials is a paradox. Judges and jurors need help with matters beyond their understanding, and judges are expected to act as gatekeepers to ensure that jurors are not fooled by misleading expert testimony. Yet, as gatekeepers, judges might not effectively distinguish sound from unsound expert testimony. As factfinders, judges and jurors both might have difficulty comprehending expert evidence, intelligently resolving conflicts between experts, and applying the scientific and technological evidence they hear to the larger dispute before them. This essay explores those problems and a variety of possible solutions, ranging from more effective ways parties might present technical information at trial, to educational interventions supervised by the court, to making juries more effective in performing their task, to more controversial measures, such as replacing conventional juries with special juries and replacing generalist judges with expert judges.

VALERIE P. HANS is Professor of Law at Cornell Law School.

MICHAEL J. SAKS is Regents Professor at Arizona State University, where he is on the faculties of the Sandra Day O'Connor College of Law and the Department of Psychology.

(Complete author biographies appear at the end of the essay.)

The fundamental paradox of the use of expert evidence in litigation is that those with the power and duty to evaluate expert testimony possess less knowledge of the specialized subject matter at issue than do the experts whose testimony they are evaluating. Judges experience this paradox not only when they are performing as factfinders in bench trials, but also when they are acting as gatekeepers of expert testimony. As one prominent judge observed:

> Though we are largely untrained in science and certainly no match for any of the witnesses whose testimony we are reviewing, it is our responsibility to . . . resolve disputes among respected, well-credentialed scientists about matters squarely within their expertise, in areas where there is no scientific consensus.[1]

The paradox also exists for juries. As Judge Learned Hand asked in 1901, "How can the jury judge between two statements each founded upon an experience confessedly foreign in kind to their own? It is

doi:10.1162/DAED_a_00527

just because they are incompetent for such a task that the expert is necessary at all."[2]

Despite this central paradox, trials by generalist judges and representative juries have much to recommend them as vehicles for the rational resolution of factual disputes involving scientific and technical issues. In other fact-finding settings, decision makers often have strong preferences or prior commitments, and even if they do not, they might be subjected to an array of pressures applied by interested parties. Consider government and industry review panels; or think about the situation facing legislators attempting to integrate into the laws they draft diverse interpretations of scientific facts pressed upon them by constituents or lobbyists. Contrast these settings with trials, where jurors and judges are expected to have no biases regarding which party prevails and what facts are found to be true. Prospective jurors are ideally to be excluded before trial if they hold beliefs or attitudes that favor one party over another, or if their own interests are linked to a side in the case. Judges are expected to recuse themselves from cases, allowing other judges to preside, if they have or might reasonably be perceived as having close ties to a party or an attorney who would appear before them or a financial or other interest in the outcome of a case.

During trials, the system uses tools for informing decision makers about relevant facts that are, by design, fundamentally concerned with guaranteeing the relevance and reliability of information. To this end, the architecture of the adversary system promises the opportunity to make counterarguments for every important claim made by an opposing advocate. Ideally, the judge and the jury hear the parties' accounts, consider the competing factual claims and interpretations urged upon them, and then do their best to reach the verdict that best fits the facts they deem most likely correct. Compared to many other settings for fact-based dispute resolution, including those involving scientific facts, courtroom trials – notwithstanding their imperfections – are among the most rationally constructed.

In trials where expert scientific evidence bears on the heart of a dispute, the key problem is not the absence of factfinder neutrality, but rather that the decision makers arrive at their task without the knowledge, and perhaps without the intellectual skills, needed to complete their assignment effectively. Thousands of trials take place in federal and state courtrooms nationwide each year, often deciding significant cases with far-reaching implications. If the trial process is to serve the parties and the larger society well, the law must find means to overcome the inherent limitations that arise when scientific expertise is needed to resolve disputes. In this essay, we offer a range of suggestions for how judge and jury fact-finding in trials with scientific evidence might be improved.

Before discussing how trials might be made to work better, it is worth illustrating challenges likely to arise. Judges have long been the gatekeepers of evidence, screening proffered testimony under rules that evolved to prevent false or misleading evidence, including expert evidence, from leading jurors astray. The admissibility decision is key: if plaintiffs cannot use scientific evidence to make their case, the case may be resolved through summary judgment or collapse on its own. Yet this arrangement applied to experts is paradoxical at its core: expert evidence must be prescreened for nonexpert jurors by nonexpert judges.

Because jurors typically (though not invariably) are laypersons lacking the expertise to evaluate scientific and other technical evidence, they are offered the guidance of experts. On occasion, courts appoint neutral expert witnesses for this purpose.[3]

But, typically, experts are provided and employed (literally) by parties who wish to lead the jurors to particular conclusions. The filter interposed to protect jurors from being misled by invalid or misleading expert testimony consists of another nonscientist, the judge, who is generally not much better situated than the jurors to decide whether what is being received is sound or not. Some judges might have the benefit of experience with similar scientific evidence or can draw on their clerks' knowledge. Nonetheless, as we suggest below, in some circumstances, judges may lack strengths jurors have in evaluating scientific evidence.

In recent years, DNA exonerations of innocent defendants have called attention to the long-standing and consequential failures of judges as gatekeepers in relation to various forensic sciences.[4] For more than a century, judges have assessed the proffered testimony of witnesses who claimed to be able to identify the source of fingerprints, bite marks, hair, handwriting, footprints, tool marks, and the like found at crime scenes. These witnesses were typically allowed to testify with little or no vetting, and they have been extraordinarily persuasive in both bench and jury trials. Consider the case of Cameron Todd Willingham.[5] The state's arson experts concluded that "arson indicators" established that a fire was intentionally set, making murders of the deaths of Willingham's children in the fire. The court admitted the expert testimony about what their so-called arson indicators implied even though there had been no empirical tests that showed that these indicators could distinguish accidental from purposefully set fires. The jurors accepted as sound the expert claims that had passed judicial muster. They convicted Willingham and sentenced him to death; he was subsequently executed.[6]

A month after Willingham's conviction, a major publication of the leading fire and arson investigation organization summarized ongoing empirical testing that found that the "indicators" relied on in the trial were unable to distinguish arson fires from accidental ones.[7] Over the next twelve years, until Willingham's execution in 2004, in the course of numerous appeals, no court was ever asked to reconsider the (in)validity of the expert testimony that had been offered at trial.[8]

Courts have rarely excluded the findings and testimony of expert forensic scientists, but in recent years, interdisciplinary bodies of scientists have reviewed those forensic offerings and declared some of them, like the arson indicators, not only to be largely or completely lacking in empirical validation, but also to be almost certainly invalid.[9] The National Research Council, the research arm of the National Academy of Sciences, which established a subcommittee to review the forensic sciences, concluded:

> The bottom line is simple: In a number of forensic science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, and the courts have been utterly ineffective in addressing this problem.[10]

Although it came too late to help Mr. Willingham, the field of fire and arson examination removed nearly two dozen "arson indicators" from its corpus of supposed knowledge when they were tested empirically and found unable to distinguish arson fires from accidental blazes.[11] Two other forensic disciplines (voiceprint identification and comparative bullet lead analysis) closed up shop after being found by scientific review bodies – but not by the courts – to lack sound bases for their claims.[12] A fourth technique, bite mark identification, seems to be next in line to be discredited, though, to date, no court has ever found it inadmissible.[13] It is unlikely to be the last forensic discipline to be shelved for failing the test of empirical validation.

Judicial gatekeepers have been unable to distinguish pseudoscience from science even after the U.S. Supreme Court, in *Daubert v. Merrell Dow Pharmaceuticals*, clarified the test of admissibility to emphasize that the touchstone for the admissibility of scientific claims is demonstrated validity.[14] These are hints that knowledge is not enough: if judges are unwilling to follow the evidence where it leads when it leads to unfamiliar destinations or unwelcome acquittals, then nonjudicial institutions will have to come to the rescue.[15] Until they become better informed about a subject, neither average judges nor average citizens are likely to have more than a limited understanding and stereotypical impressions of the multitude of scientific and technical fields, and little ability to critically evaluate those fields' claims.[16]

When judges decide to admit scientific evidence, they risk putting an unintended thumb on the scale. Consider psychologists N. J. Schweitzer and Saks's research finding evidence of a "gatekeeper effect."[17] Participants evaluated expert evidence presented within or outside of a trial context. Those who reviewed evidence they believed had successfully passed through a judicial filter regarded the evidence as being of higher quality and more persuasive than participants who evaluated evidence presented outside the trial context. Apparently, participants assumed that evidence that survives the law's seemingly rigorous gatekeeping can be regarded as sound science.

Of course, even if a judge conscientiously and correctly admits only acceptably sound science, problems can remain, for some scientific issues are legitimately disputed between equally knowledgeable and sincere experts. How is the jury to referee such a dispute? Making matters even more difficult, because the great majority of cases are disposed of before trial, and because pretrial settlement tends to remove the clearest and easiest cases, what lands in court are the cases that the parties and their lawyers were unable to resolve, sometimes because of profound disputes over the facts. Thus, what the legal process delivers to judges and juries tends to be the most unclear, ambiguous, and challenging of the mass of cases initially filed.

Research indicates that when people are motivated and able to do so, they engage in central, or "System 2," processing: that is, they process information thoughtfully in an effort to solve the problem confronting them.[18] But when they are unmotivated or unable, perhaps due to lack of ability or information overload, they tend to engage in peripheral, or "System 1," processing, relying on superficial features of the information before them, such as the number of arguments or the characteristics of the witnesses and attorneys.[19] Thus, in trials in which jurors (or judges) might be overwhelmed by unfamiliar scientific evidence and confused or frustrated by testimony beyond their comprehension, shallow System 1 thinking may seriously endanger sound fact-finding.

Recent research suggests that even when expert testimony is presented in a relatively straightforward fashion, laypeople may be insensitive to the empirical support for a proposition (or lack thereof), although scientists see empirical tests as the touchstone for resolving scientific disputes. Instead, they may rely more on the background and experience of the witness presenting the evidence as a measure of the testimony's value. Although credentials can be informative, lawyers for both parties may seek out and succeed in hiring expert witnesses with similarly impressive credentials. If they do, the evaluation of experts' credentials will supply an even less reliable means of determining which opposing expert is the more competent.

Quantitative, statistical, and probability evidence can be especially confusing and potentially misleading. For example, stu-

dents in a college economics class were upset because their grades averaged 72 (out of a maximum 100 points), even though they were graded on a curve and the distribution of A's, B's, and C's was a predetermined constant. On the next exam, the professor employed a raw scale maximum of 137, on which the average score was now 96 (actually implying *poorer* performance by the class as a whole). Again, earned grades reflected the students' relative position in the class and the same number of A's, B's, and C's were given as before. This time, however, the students were much happier. A class average of 96 felt better than 72. The students were influenced emotionally by the superficial impression made by the raw scores, even though they understood cognitively that what mattered was their relative rank on the raw scale, whatever the scale happened to be.[20]

Jurors try to fit the evidence they hear into stories, narrative accounts that make sense of the facts of a case and imply particular case outcomes. Like most of us, they struggle hard to understand statistical and probability evidence and to infer its implications for a case. Typically, people underutilize such evidence in their decision-making and are more influenced by clinical evidence than they are by more diagnostic actuarial evidence.[21] In trials, there is reason to think the problem is especially acute. Expert evidence, especially of the statistical kind, is difficult to incorporate into a story of the case, thus inviting undervaluation in comparison with other, more case-specific, narrative kinds of testimony.

Even when people understand the relevance of probability evidence, they can make "misaggregation errors," causing them to underutilize the evidence. A misaggregation error occurs "when a person's subjective belief in the validity of a hypothesis [e.g., the defendant is guilty] is not updated to the extent that is logically warranted based on prior beliefs and the probative value of a new piece of probabilistic evidence."[22] Relatedly, people underadjust for laboratory error rates when assessing the meaning of a forensic test's results.[23]

In some contexts, however, probability data can be overweighted. A well-known example is the "prosecutor's fallacy," which confuses the frequency of a trait in the population (for example, one person in a million has DNA that matches the crime scene DNA) with the probability that someone *other than* the defendant left the evidence showing that trait (that there is only one chance in a million that crime scene DNA came from someone other than the defendant). Further illustrating the confusion that probabilistic evidence can cause, if the same data are presented as frequencies rather than as probabilities (such as one out of every million people has DNA that would match the crime scene DNA), this can produce the opposite effect: undervaluing the probative value of the evidence given the other evidence in the case.[24]

Civil cases present another broad range of challenges for factfinders.[25] Jurors and judges alike can easily become confused by material presented during expert testimony in civil trials, such as the meaning of statistical significance, practical significance, confidence intervals, relative versus absolute risk, or regression models.

In addition to the difficulties of dealing with statistics, most if not all of the heuristics and biases made famous by psychologists Daniel Kahneman and Amos Tversky can foster distortions in the rational interpretation of information and lead to error. Even experts are susceptible to such sources of error. For example, physicians who regularly counsel patients on the results of screening tests like mammograms sometimes make erroneous inferences about the meaning of a positive test result, even when they have all the information needed to reach a correct interpretation.[26]

Many of the problems of comprehending, evaluating, and using unfamiliar technical evidence to make important decisions are not peculiar to jurors. They are problems for most people in most situations, certainly including judges, and sometimes or often including trained specialists, who should have a fighting chance to get things right, but who are incompletely schooled in the evidence or fall prey to misleading cognitive heuristics.

The situation is not, however, entirely bleak. Even in cases with extensive scientific evidence, some factual disputes do not demand expert analysis. Instead, their resolution turns on credibility or related judgments. These may reflect not just the credentials of rival experts but the consistency of their claims, or the way witnesses hold up under cross-examination, or judgments about facts in dispute that the experts mutually acknowledge to be dispositive. In the medical malpractice area, for example, various studies have assessed the reasonableness of jury verdicts. Some studies have compared jury verdicts with confidential assessments of the same cases made by neutral physicians. These studies generally find agreement between the physicians and the juries.[27]

Even when testimonial or other evidence is unfamiliar and complex, jurors and judges can absorb and ponder the evidence deeply (central processing), even if mixed with other, more superficial thinking (peripheral processing).[28] Thus, a reasonable goal for improving the use of expert evidence is to find ways to facilitate an increased ratio of central to peripheral processing of trial information.[29]

Trials offer fact-finding benefits as well as challenges. The advantages might be leveraged for further improvement. We offer specific suggestions below, some modest, others more controversial. Some are based on findings derived from empirical research; others are in need of testing. These include:

- Presenting expert evidence to maximize understanding;
- Restructuring the trial to maximize understanding;
- Implementing trial procedure reforms that promote understanding;
- Educating judges;
- Educating juries;
- Ensuring diverse juries and robust deliberation; and
- Changing the factfinder to special juries and expert judges.

Trials are inherently educational forums. The whole exercise is about communicating relevant information to factfinders for decision-making. Trial procedures can be tweaked so that their capacity for educating is improved. Judges have considerable discretion to manage evidence before and during the trial, so long as they do not unduly burden the fundamental right of the parties to assemble and present the evidence.

Where there is a battle of experts, jurors may end up skeptical of both sides, undermining their use of relevant expert evidence in the decisions.[30] But smart and capable lawyers on one or both sides, with the cooperation of the judge, should be able to find ways to work with their experts to provide factfinders with sound and comprehensible information, such that the case rooted in sounder science helps itself while facilitating better decision-making.

There are ways to present unfamiliar or complex information so that it can be better understood and used in trial decisions. Attorneys and their expert witnesses can and should adopt these methods. Psychologist Gerd Gigerenzer and colleagues, for example, have put much energy into finding ways to make statistical presentations more intuitively understandable. Their suggestions include: use numbers, not just words to de-

scribe quantities and risks; present numbers in data tables; use natural frequencies rather than conditional probabilities; use frequencies rather than single-event probability statements; and report absolute risks, not relative risks.[31] Researchers have also recommended communicating numerical information using visual aids such as bar graphs, pie charts, 2x2 tables, and Venn diagrams. Well-crafted visual displays can help jurors understand probabilities and magnitudes and can help them avoid framing effects (a form of cognitive bias resulting from how information or questions are presented).[32]

Applying these and other educational approaches to the courtroom context has thus far generated mixed results. In one study, participants from a county jury pool had great difficulty inferring causality from the data in a 2x2 contingency table representing evidence in a toxic tort case (one in which the claim is that exposure to a toxic substance caused some person or persons to suffer adverse health effects). None of the various explications by an epidemiologist expert witness (how contingency tables work, how relative risk and odds ratios are calculated, how to properly interpret contingency table data) improved the participants' ability to reach correct inferences about causation or its absence.[33] More testing of suggested techniques is needed in trial settings, but we are optimistic that research will find ways that enable attorneys and their expert witnesses to make more comprehensible the evidence they present to juries. Courts might also consider sharing experts' reports with juries, whether or not the parties request it.

In addition to clearer presentations, experts could help jurors by conveying more about areas of consensus in their fields. Some experts are criticized for advocating idiosyncratic views at odds with the majority view in their field, but judges and jurors without specialist knowledge have little ability to determine how common or infrequent the allegedly idiosyncratic views are. Though being in the mainstream is no guarantee of correctness, survey studies of experts about where the consensus lies regarding various phenomena could help factfinders put a trial expert's assertions in context.[34]

Judges have more power to regulate trial structure and proceedings than they typically exercise. Before a trial begins, courts could work harder with the parties to help them resolve disputes and stipulate to the conclusions to some, if not all, of the highly technical issues that might arise in a case, thereby removing them from controversy at trial (with jurors instructed on what the agreed-upon conclusions were). "Hot-tubbing," a procedure used in Australia and Canada that begins with experts meeting together without the parties or their lawyers before the trial, could aid in identifying areas of agreement and disagreement, as Nancy Gertner and Joseph Sanders discuss in their contribution to this issue.[35]

At trial, judges might improve their own as well as jurors' comprehension by requiring that opposing evidence on difficult scientific or technical issues be offered back-to-back, juxtaposing expert witnesses with competing views on the same topic.[36] Thus, instead of hearing from a plaintiff's expert witness and not hearing from the defense's rebuttal expert until much later, a court could order that the direct and cross-examination of the defense witness occur immediately following the direct and cross of the plaintiff's expert. This procedural reform, however, is not without its challenges, as Gertner and Sanders discuss in greater depth.

On the criminal side, resources are often so imbalanced that special funding or other procedures are needed so that the defense is able to present expert evidence and prevent gatekeeping judges and factfind-

ers from reaching decisions based on incomplete or distorted pictures of the state of the science. Our discussion of forensic sciences that lack scientific validity and the large role that forensic science evidence has played in producing wrongful convictions provides more than enough cautionary tales to justify wariness about one-sided presentations by interested experts.[37]

Optimally educating jurors will require changes in the way courts do things. If trials are to serve the parties and the larger society, means must be found to overcome inherent limitations that exist at the outset of the trial process. Traditional jury trials operate with the assumption that jurors are empty vessels who passively receive the evidence presented by the parties, refrain from forming even preliminary opinions, and wait until the trial has concluded to deliberate and decide the case. In most courtrooms, jurors are not allowed to ask questions of the witnesses or to talk with one another until the end of the trial. In a traditionally conducted complex trial, juror confusion and mistakes in interpreting scientific testimony during the case presentation can neither be detected nor corrected as they occur.

The American Bar Association's 2005 report *Principles for Juries and Jury Trials* advocates "active jury" trial practices to promote juror understanding.[38] Allowing jurors to clarify evidence and issues by permitting them, under carefully controlled conditions, to submit questions for witnesses, and allowing jurors to talk to one another during the trial so they can discuss scientific evidence while it is fresh in their minds, could promote better understanding and use of scientific evidence.

There is now a modest body of research on active jury reforms, including note-taking, question-asking, and juror discussions. Jurors who serve in trials in which they are able to ask questions and to talk with other jurors during breaks have provided generally positive feedback about these changes, and few if any negative effects have been detected.[39] Jurors who have the opportunity to take notes also typically perform better.[40] One experiment assessing how well mock jurors understood scientific evidence found that those using checklists and jury notebooks performed better than jurors not allowed to employ these innovations.[41]

Judges and lawyers often greet active jury reforms with skepticism, but most change their views after participating in a trial in which the reforms are employed. The scientists and engineers surveyed by Shari Diamond and Richard Lempert and reported on in this issue also appear to prefer a more educational approach: 57.7 percent said they would be more likely to participate as expert witnesses if they could answer jurors' questions following their testimony.[42]

Judges have a variety of educational programs in law and science from which to choose. These range from panels lasting an hour or two in continuing judicial education programs, to day-long focused sessions, to four-to-six-week summer courses at universities such as Duke and Virginia. The potentially most useful of these efforts seek to teach judges how to be more thoughtful, critical consumers of specialized knowledge. We are not, however, aware of any systematic empirical attempts to see whether these efforts have enabled judges to better understand the science and technology issues that arise when they preside over trials.

Other programs focus on substantive science. For example, the Federal Judicial Center's (FJC) Education Division collaborates with universities to offer short courses on such topics as neuroscience and law, law and the biosciences, and the economics of antitrust law.[43] Former Education Di-

vision Director Bruce Clark told us that judicial education programs at the FJC and elsewhere are increasingly using more active, engaged methods of teaching, which seems promising.

The *Reference Manual on Scientific Evidence*, now in its third edition, also attempts to educate judges on science. It provides well-informed guides to specific scientific fields, written by experts in those fields.[44] The goal is to aid judges in managing cases with scientific and technical evidence. Chapters review and explain the science that commonly arises in legal cases, including such matters as DNA analysis, engineering, mental health evidence, survey methodology, epidemiology, and statistics.

Researchers have also suggested tutorials on technical and scientific topics for judges. Litigators Jeffrey Snow and Andrea Reed have outlined an approach to using tutorials to educate judges in patent cases: "The technical tutorial has few common ground rules. In its most general form, the technical tutorial is a non-evidentiary presentation for the educational benefit of the district court judge."[45] They distinguish between an adversarial approach to construction of the tutorial, in which each party has its own experts explain the underlying science, and the possibility of having both parties agree on a neutral court-appointed expert to provide technical background as a witness. Alternatively, the parties might collaborate on a report or video that the judge can review on his or her own. Although tutorials seem useful and judges request them, we know of no research on the effectiveness of technical tutorials in patent cases.

Jurors, too, might receive pretrial education and training through tutorials tailored to the science they are likely to encounter. However, although the idea has been floated and used on at least a few occasions, we know of no jurisdiction where it has been implemented as a routine practice when scientific evidence is involved.[46]

Research suggests that brief yet effective education in specific intellectual skills is possible. Social psychologist Richard Nisbett and colleagues have developed and tested a training intervention that attempts to teach laypeople the statistical concept of the "law of large numbers." The intervention consists of two parts: "rule training" involves reading a description of the law of large numbers, and "example training" involves a worksheet containing three sample problems that highlight the various principles of the law of large numbers, followed by a written explanation and analysis of the problems. The greatest improvement in statistical reasoning was achieved by those participants who received *both* rule- and example-based training.[47]

Using the rule-plus-example approach, Schweitzer and Saks tried to improve upon past (unsuccessful) efforts to train jurors to understand scientific causation.[48] The brief, non–case-specific intervention aimed to teach jurors to understand and identify the three requisites of causal inference: temporal precedence, covariation, and nonspuriousness. Jurors' grasp of the concepts was tested by presenting a videotaped mock toxic tort trial. The critical evidence was a study, presented by an expert witness, that tested the causal relationship between the defendant's product and lung disease through either a properly designed experiment or one in which one or another of the key elements of causal inference was absent. Untrained jurors were unable to distinguish the well-designed experiment from any of the defectively designed experiments. Trained jurors were better able to assess the quality of the research, and their verdicts reflected their sounder understanding.

Jonathan Koehler would go further and provide jurors with a "comprehensive pretrial training program" that would teach logical inference, how to distinguish be-

tween weak and strong evidence, how to combine pieces of evidence, and how to apply law to facts; test the jurors' performance; and exclude from service those who are not up to par.[49] Excluding jurors on these grounds might well undermine the jury's ability to represent the community, however.

Perhaps the most ambitious study to date of jury tutorials is an Australian project that gave only some mock jurors hearing a DNA case a DNA tutorial as part of the expert evidence in a case.[50] The tutorial, developed in consultation with scientific and forensic experts, devoted twelve minutes to the science of DNA profiling and five minutes to understanding random match probabilities, a key concept in assessing the meaning of a DNA match. Some participants heard an expert orally deliver the tutorial, while others heard an expert give the same talk accompanied by multimedia displays. Still others served in a control condition, receiving no expert evidence. Mock jurors then decided a case in which the DNA evidence was crucial. Most participants began knowing little about DNA. Those who started knowing the least about DNA knowledge tended to express undue belief in DNA evidence; those knowing more about DNA were more skeptical at the start of the trial. The expert evidence that included the DNA tutorial significantly improved jurors' understanding. Compared with those in the control condition, who received no tutorial, those hearing any version of the tutorial showed greater comprehension of DNA identification.

In this study, the multimedia presentation of evidence did not significantly improve comprehension beyond the gains produced by the oral presentation alone, though it did more to close the gap between less knowledgeable jurors and those with greater knowledge. Whether the same would be true of such dramatically new media forms as virtual reality and augmented reality cannot be known, but these applications might turn out to be unusually effective and efficient teaching tools.[51]

The FJC has developed tutorials for use in patent jury trials.[52] Roderick McKelvie, then a district court judge in Delaware, encouraged the FJC to prepare a tutorial video to educate juries in patent trials. He joined a group of patent lawyers and judges who contributed to the text for the video, which was then reviewed by the FJC and other experts. The first video, seventeen minutes long, was released in 2002 and updated in 2013. The videos did not seek to educate jurors on the scientific matters at issue in a case, but rather offered background information about what a patent is, the place of patents in society, and the work of the U.S. Patent and Trademark Office (PTO). The FJC aimed "to present a balanced view of the patent process" but cautioned judges to "review it carefully and consult with counsel before deciding whether to use it in a particular case."[53]

Some patent lawyers criticized the 2002 video as unbalanced.[54] The script did not concern them, but the images did. The visual portrayal of "conscientious, hard-working examiners" seemed to favor patentees, although other images of the "piles" of patent applications and "endless rows" of files seemed to suggest overworked and overwhelmed patent examiners, favoring defendants. One jury-consulting firm presented the 2002 video in mock jury exercises in five venues across the United States.[55] Mock jurors' responses before and after seeing the tutorial were compared, showing dramatic improvements in reported understanding of patents. For example, before watching the video, a majority (57 percent) said they did not understand what a patent claim was, but that number dropped to 4 percent after the video. Just 24 percent initially knew that a patent granted by the PTO could be invalidated by a judge or a jury; afterward, that number jumped to 63 percent. The consultant con-

cluded that the video was effective in educating juries about both pro-plaintiff and pro-defense perspectives. A repeat of the study using the 2013 FJC patent video produced similar results.[56] Research that examines whether patent tutorials improve juror understanding of expert evidence in patent trials would be of substantial value.

The fact that juries engage in group decision-making allows juries to bring more intellectual resources to their task than any one person, including a judge, can deliver. Indeed, juries have the potential, depending upon the methods used to recruit them, to possess knowledge, experience, and analytic capacity that exceeds that of most judges. The sheer fact that juries are groups provides advantages. Where all citizens are required to serve, with very limited excuses granted, juries will be composed of people from all kinds of educational and occupational backgrounds. This means they will not infrequently include people with scientific, technical, and quantitative capabilities that few judges possess.[57]

The more jurors on the jury, the greater the chances of having some who are able to understand difficult subject matter. If the trend toward smaller juries of six or eight cannot be reversed entirely, complex cases at least ought to be tried to twelve jurors because deliberations are likely to be richer with greater educational potential. An individual juror who has a better grasp of the scientific evidence presented at a trial can explain the meaning and significance of the evidence to the other jurors, increasing their ability to properly weigh the scientific information.[58]

In a mock jury experiment in which mitochondrial DNA (mtDNA) was the focus of expert testimony, researchers examined the impact of deliberation on jurors with lower and higher levels of comprehension.[59] Jurors' prior knowledge, as evidenced by science and mathematics courses they had taken, increased their ability to benefit from deliberation. However, mock jurors with lower initial levels of comprehension gained the most from deliberations.

Judges (or special masters appointed by judges to initially hear cases and report back on their findings) are sometimes suggested as an alternative to the jury in complex cases. Several decades ago, there were cases in which lawyers asked courts to recognize a "complexity exception" to the right to a jury trial, arguing that where it was thought that juries could not adequately understand the evidence, the case must be tried to a judge. Appellate courts divided on whether such an exception should exist.[60] Regardless of whether a party might be denied a jury, it is worth noting that generalist judges may be no more able to master the intricacies of complex, expert scientific testimony than a representative jury. Reviewing a set of complex cases, Lempert concluded that when judges were competent and well organized, the juries they supervised were effective as well.[61] If judges are to be used as an alternative to juries, they might do better if drawn from panels specially chosen for having relevant knowledge or if they sat as three-judge courts.

Complex cases might also be tried by special juries, drawn from pools of people with more formal education or particularly relevant experience or training. Special or "blue ribbon" juries have a long history in England and the United States. The earliest documented special jury convened in England in 1351: a jury of cooks and fishmongers for a defendant charged with selling bad food.[62] Other special juries in early England included juries of matrons tasked with determining whether a woman defendant was with child, and a jury of businesspeople in a business contract case. In the United States, there was a time when almost half the states had special jury statutes for use in cases of high importance

or great difficulty, although that number has dwindled. Special juries in the United States are rare today, owing partly to statutory requirements of cross-sectional representation on jury panels, but also to increased appreciation of the fact-finding benefits and symbolic significance of representative juries.[63] Even without using a special jury, judges and lawyers could employ *voir dire* questions to explore scientific competence in an effort to increase the proportion of highly numerate or better-educated jurors.

Numerous studies have found that people with higher educational attainment generally and greater familiarity with mathematics and science in particular are better able to understand scientific and other technical information and to apply that understanding to solving problems.[64] People high in numeracy have been found better able than their low-numeracy peers to comprehend and apply numerical principles, and they are somewhat less susceptible to being influenced by framing and other irrelevant factors.[65] Research on the dynamics of juries with one or a few such members is limited. But, clearly, the juries they are on have the potential to benefit from their more knowledgeable members. Some studies have found that jurors with relevant knowledge are recognized by their peers and placed in leadership positions.[66] To what extent their oversized influence is beneficial or not remains to be discovered.

We raise three caveats about these special juries. First, it is clear that numeracy and advanced education are not panaceas.[67] Judges and highly numerate individuals make processing mistakes and are influenced by common heuristics and biases.[68] Second, recent research finds that in controversial areas of science, people with substantial backgrounds and advanced education in a field may be more biased in their evaluations than those who are less knowledgeable.[69] Relatedly, these highly knowledgeable jurors tend to be disproportionately influential in the jury deliberation, as others defer to their superior knowledge. Third, selecting jurors using one attractive characteristic may have unexpected negative consequences, since individual characteristics do not exist in isolation. More men than women major in science, for example. Educational attainment is linked to race, income, and political affiliation. Blue ribbon juries are likely to fail to adequately reflect the attitudes and experiences of the community, particularly in deciding on matters like damages. Moreover, scientific matters may not be the only matters in dispute; correctly resolving a purely scientific question may be only one part of the decision. As we discussed above, diverse juries composed of people from different parts of the community have their own fact-finding advantages, which could be lost if we selected jurors mainly for their educational attainment.

Generalist judges and lay juries face considerable challenges in trials with scientific evidence. Yet the adversary trial provides us with opportunities to modify procedures or educate or select factfinders to maximize the ability of judges and juries to understand expert scientific evidence and to use it effectively to resolve a case. We have suggested a number of reforms, but more study of possible changes is needed. We must collect data and run experiments; that is, we should take a scientific approach to deciding on those reforms that will best enable judges and juries to cope with modern scientific evidence.

AUTHOR BIOGRAPHIES

VALERIE P. HANS is Professor of Law at Cornell Law School. She is the author of *The Psychology of Tort Law* (with Jennifer K. Robbennolt, 2016), *American Juries: The Verdict* (with Neil Vidmar, 2007), and *Business on Trial: The Civil Jury and Corporate Responsibility* (2000).

MICHAEL J. SAKS is Regents Professor at Arizona State University, where he is on the faculties of the Sandra Day O'Connor College of Law and the Department of Psychology. He is the author of *The Psychological Foundations of Evidence Law* (with Barbara A. Spellman, 2016) and author/editor of *Modern Scientific Evidence: The Law and Science of Expert Testimony* (with David L. Faigman, Joseph Sanders, and Edward K. Cheng, 2009 – 2010).

ENDNOTES

[1] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 43 F.3d 1311, 1316 (1995). The common law allowed expert evidence to be introduced only when the expert was testifying to matters beyond the ken of the average juror. This requirement was incorporated into Federal Evidence Rule 702, which allowed for expert evidence in circumstances in which the knowledge would assist the trier of fact in understanding the evidence or determining a fact in issue.

[2] Learned Hand, "Historical and Practical Considerations Regarding Expert Testimony," *Harvard Law Review* 15 (1) (1901): 54.

[3] Daniel L. Rubinfeld and Joe S. Cecil, "Scientists as Experts Serving the Court," *Dædalus* 147 (4) (Fall 2018).

[4] Brandon Garrett, *Convicting the Innocent: Where Criminal Prosecutions Go Wrong* (Cambridge, Mass.: Harvard University Press, 2012).

[5] John Lentini, "Fires, Arsons, and Explosions," in *Modern Scientific Evidence: The Law and Science of Expert Testimony*, ed. David L. Faigman, David H. Kaye, Michael J. Saks, et al. (Toronto: Thomson Reuters, 2010).

[6] Steve Mills and Maurice Possley, "Texas Man Executed on Disproved Forensics," *Chicago Tribune*, December 9, 2004 [describing Willingham case].

[7] National Fire Protection Association, *NFPA 921: Guide for Fire and Explosion Investigations* (Quincy, Mass.: National Fire Protection Association, 1st ed., 1992; 2nd ed., 1995).

[8] Rachel Dioso-Villa, "Scientific and Legal Developments in Fire and Arson Investigation Expertise in *Texas v. Willingham*," *Minnesota Journal of Law, Science, and Technology* 14 (2013).

[9] National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (Washington, D.C.: National Academies Press, 2009); and President's Council of Advisors on Science and Technology, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* (Washington, D.C.: Executive Office of the President, 2016).

[10] National Research Council, *Strengthening Forensic Science in the United States*, 1 – 14 [see note 9].

[11] John Lentini, "Fires, Arsons, and Explosions," in Faigman et al., eds., *Modern Scientific Evidence* [see note 5].

[12] National Research Council, *On the Theory and Practice of Voice Identification* (Washington, D.C.: National Academies Press, 1979); National Research Council, *Committee on Scientific Assessment of Bullet Lead Elemental Composition Comparison, Forensic Analysis: Weighing Bullet Lead Evidence* (Washington, D.C.: National Academies Press, 2004); and Paul C. Giannelli, "Forensic Science: *Daubert*'s Failure," *Case Western Reserve Law Review* (forthcoming), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3031227.

[13] Michael J. Saks, Thomas Albright, Thomas L. Bohan, et al., "Forensic Bitemark Identification: Weak Foundations, Exaggerated Claims," *Journal of Law and the Biosciences* 3 (2016).

[14] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

[15] When he was a federal district judge deciding *Daubert* on remand from the U.S. Supreme Court, Alex Kozinski (later, chief judge of the U.S. Ninth Circuit Court of Appeals) suggested that forensic science should be exempted from the validity test that he was required to apply to the epidemiological analysis he found inadequate in *Daubert*. At that time, he believed unreservedly in the claims of forensic science experts, without expecting them to demonstrate empirically the soundness of their assertions. *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 43 F.3d 1311 (1995). Kozinski, after criticizing experts who acquire their knowledge for purposes of litigation, states that "there are, of course, exceptions. Fingerprint analysis, voice recognition, DNA fingerprinting and a variety of other scientific endeavors closely tied to law enforcement may indeed have the courtroom as a principal theatre of operations. . . . As to such disciplines, the fact that the expert has developed an expertise principally for purposes of litigation will obviously not be a substantial consideration." Two decades later, after working with a White House advisory council on the forensic sciences, his views had changed dramatically: "Virtually all of these methods are flawed, some irredeemably so." "Why," he asked, "trust a justice system that imprisons and even executes people based on junk science?" See Alex Kozinski, "Rejecting Voodoo Science in the Courtroom," *The Wall Street Journal*, September 19, 2016.

[16] Richard Posner and Frank Easterbrook, who at the time were both eminent judges of the U.S. Seventh Circuit Court of Appeals, famously excoriated both judges and lawyers for their fear of science. *Jackson v. Pollion*, No. 12-2682 (7th Cir. 2013).

[17] N. J. Schweitzer and Michael J. Saks, "The Gatekeeper Effect: The Impact of Judges' Admissibility Decisions on the Persuasiveness of Expert Testimony," *Psychology, Public Policy, and Law* 15 (1) (2009): 1 – 18.

[18] Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus and Giroux, 2011).

[19] See, for example, Special Committee on Jury Comprehension of the American Bar Association Section of Litigation, *Jury Comprehension in Complex Cases* (Chicago: American Bar Association, 1989); and Joseph Sanders, Shari S. Diamond, and Neil Vidmar, "Trial Lawyer Perceptions of Expert Knowledge," *Psychology, Public Policy, and Law* 8 (2) (2002): 139 – 153. Richard Lempert, "Civil Juries and Complex Cases: Taking Stock After Twelve Years," in *Verdict: Assessing the Civil Jury System*, ed. Robert E. Litan (Washington, D.C.: Brookings Institution Press, 1993), finds that jury performance is more likely to be a problem in those cases that are "complex" because of the technical nature of the evidence; and Joseph Sanders, "The Jury Deliberation in a Complex Case: *Havner v. Merrell Dow Pharmaceuticals*," *Justice System Journal* 16 (2) (1993): 45 – 67, reaches a similar conclusion based on jurors' answers to researcher questions about testimony in a complex evidence case involving the drug Bendectin, concluding that the jurors had a weak grasp of the science resulting in an indefensible verdict.

[20] Richard H. Thaler, *Misbehaving: The Making of Behavioral Economics* (New York: W. W. Norton and Company, 2015).

[21] Joseph Sanders, Michael J. Saks, and N. J. Schweitzer, "Trial Factfinders and Expert Evidence," in Faigman et al., eds., *Modern Scientific Evidence* [see note 5]; Daniel A. Krauss and Bruce D. Sales, "The Effects of Clinical and Scientific Expert Testimony on Juror Decision Making in Capital Sentencing," *Psychology, Public Policy, and Law* 7 (2) (2002): 267 – 310; and Shari Seidman Diamond, Jonathan D. Casper, Cami L. Heiert, and Anna-Maria Marshall, "Juror Reactions to Attorneys at Trial," *Journal of Criminal Law and Criminology* 87 (1) (1996): 17 – 47.

[22] Jason Schklar and Shari Seidman Diamond, "Juror Reactions to DNA Evidence: Errors and Expectancies," *Law and Human Behavior* 23 (2) (1999): 159 – 184.

[23] Jonathan J. Koehler, "When Are People Persuaded by DNA Match Statistics?" *Law and Human Behavior* 25 (5) (2001): 493 – 513; and William C. Thompson, "Are Juries Competent to Evaluate Statistical Evidence?" *Law & Contemporary Problems* 52 (1989): 9 – 41.

[24] Koehler, "When Are People Persuaded by DNA Match Statistics?" [see note 23].

[25] Richard Lempert, "Befuddled Judges: Statistical Evidence in Title VII Cases," in *Legacies of the 1964 Civil Rights Act*, ed. Bernard Grofman (Charlottesville: University of Virginia Press, 2000).

[26] Gerd Gigerenzer, *Calculated Risks: How to Know When Numbers Deceive You* (New York: Simon and Schuster, 2002).

[27] Neil Vidmar, "Are Juries Competent to Decide Liability in Tort Cases Involving Scientific/Medical Issues? Some Data from Medical Malpractice," *Emory Law Journal* 43 (1994): 885–911; and Neil Vidmar, *Medical Malpractice and the American Jury: Confronting the Myths about Jury Incompetence, Deep Pockets, and Outrageous Damage Awards* (Ann Arbor: University of Michigan Press, 1997).

[28] Shari Seidman Diamond and Jonathan D. Casper, "Blindfolding the Jury to Verdict Consequences: Damages, Experts and the Civil Jury," *Law and Society Review* 26 (3) (1992): 513–564.

[29] For example, see Neil J. Vidmar and Regina A. Schuller, "Juries and Expert Evidence: Social Framework Testimony," *Law and Contemporary Problems* 52 (4) (1989): 133–176.

[30] Diamond and Casper, "Blindfolding the Jury to Verdict Consequences" [see note 28]; Mitchell Pacelle, "Contaminated Verdict," *The American Lawyer*, December 1986, 75–80; Molly Selvin and Larry Pinkus, *The Debate over Jury Performance: Observations from a Recent Asbestos Case* (Santa Monica, Calif.: The RAND Corporation, 1987); and Daniel Shuman and Anthony Champagne, "Removing the People from the Legal Process: The Rhetoric and Research on Judicial Selection and Juries," *Psychology, Public Policy, and Law* 3 (2–3) (1997): 242–258.

[31] Gerd Gigerenzer, Wolfgang Gaissmaier, Elke Kurz-Milcke, Lisa M. Schwartz, and Steven Woloshin, "Helping Doctors and Patients Make Sense of Health Statistics," *Psychological Science in the Public Interest* 8 (2) (2008): 53–96; and Gerd Gigerenzer and Adrian Edwards, "Simple Tools for Understanding Risks: From Innumeracy to Insight," *British Medical Journal* 327 (7417) (2003): 741–744.

[32] Rebecca Helm, Valerie P. Hans, and Valerie F. Reyna, "Trial by Numbers," *Cornell Journal of Law and Public Policy* 27 (107) (2017): 107–143.

[33] Molly Treadway, "An Investigation of Juror Comprehension of Statistical Proof of Causation" (unpublished Ph.D. diss., John Hopkins University, 1990), 54–58.

[34] Saul M. Kassin, V. A. Tubb, H. M. Hosch, and A. Memon, "On the 'General Acceptance' of Eyewitness Testimony Research: A New Survey of Experts," *American Psychologist* 56 (5) (2001): 405–416. For further discussion, see Shari Seidman Diamond, "Reference Guide on Survey Research," in Stephen G. Breyer, Margaret A. Berger, David Goodstein, et al., *Reference Manual on Scientific Evidence*, 3rd ed. (Washington, D.C.: Federal Judicial Center and National Academy of Sciences, 2011), 359–423.

[35] Edie Greene and Natalie Gordon, "Can the 'Hot Tub' Enhance Jurors' Understanding and Use of Expert Testimony?" *Wyoming Law Review* 16 (2) (2016): 359–386. See also Nancy Gertner and Joseph Sanders, "Alternatives to Traditional Adversary Methods of Presenting Scientific Expertise in the Legal System," *Dædalus* 147 (4) (Fall 2018).

[36] G. Thomas Munsterman, Paula L. Hannaford-Agor, and G. Marc Whitehead, *Jury Trial Innovations*, 2nd ed. (Williamsburg, Va.: National Center for State Courts, 2006). See also Gertner and Sanders, "Alternatives to Traditional Adversary Methods of Presenting Scientific Expertise in the Legal System" [see note 35].

[37] Michael J. Saks and Jonathan J. Koehler, "The Coming Paradigm Shift in Forensic Identification Science," *Science* 309 (5736) (2005): 892–895.

[38] American Bar Association, *Principles for Juries and Jury Trials* (Chicago: American Bar Association, 2005); and Valerie P. Hans, "U.S. Jury Reform: The Active Jury and the Adversarial Ideal," *St. Louis University Public Law Review* 21 (1) (2002): 85–97.

[39] Hans, "U.S. Jury Reform" [see note 38]; and Valerie P. Hans, David H. Kaye, B. Michael Dann, Erin J. Farley, and Stephanie Albertson, "Science in the Jury Box: Jurors' Comprehension of Mitochondrial DNA Evidence," *Law and Human Behavior* 35 (2011): 60–71.

[40] B. Michael Dann and Valerie P. Hans, "Recent Evaluative Research on Jury Trial Innovations," *Court Review* 41 (2004): 12–19 [reviewing studies of juror notetaking]; and Lynne ForsterLee

and Irwin Horowitz, "The Effects of Jury-Aid Innovations on Juror Performance in Complex Civil Trials," *Judicature* 86 (4) (2003): 184 – 189 [research study showing improved performance for jurors allowed to take notes].

41 B. Michael Dann, Valerie P. Hans, and David H. Kaye, "Can Jury Trial Innovations Improve Juror Understanding of DNA Evidence?" *Judicature* 90 (2007): 152 – 156.

42 Shari Seidman Diamond and Richard O. Lempert, "When Law Calls, Does Science Answer? A Survey of Distinguished Scientists and Engineers," *Dædalus* 147 (4) (Fall 2018).

43 Federal Judicial Center, "Programs and Resources for Judges, Special Focus Programs," https://www.fjc.gov/education/programs-and-resources-judges#SF (accessed May 26, 2017).

44 Breyer et al., *Reference Manual on Scientific Evidence*, 3rd ed [see note 34].

45 Jeffrey L. Snow and Andrea B. Reed, "Technical Advisors and Tutorials: Educating Judges," *Intellectual Property Litigation* 21 (1) (2009): 1, 21 – 22.

46 Munsterman et al., *Jury Trial Innovations* [see note 36].

47 Richard E. Nisbett, Geoffrey T. Fong, Darrin R. Lehman, and Patricia W. Cheng, "Teaching Reasoning," *Science* 238 (1987): 625 – 631; and Geoffrey T. Fong, David H. Krantz, and Richard E. Nisbett, "The Effects of Statistical Training on Thinking about Everyday Problems," *Cognitive Psychology* 18 (1986): 253 – 292.

48 N. J. Schweitzer and Michael J. Saks, "Jurors and Scientific Causation: What Don't They Know, and What Can Be Done About It?" *Jurimetrics Journal* 52 (2012): 433 – 455.

49 Jonathan J. Koehler, "Train Our Jurors," Northwestern University School of Law Faculty Working Paper 141 (Chicago: Northwestern University School of Law, 2006), http://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=1140&context=facultyworkingpapers.

50 Jane Goodman-Delahunty and Lindsay Hewston, "Improving Jury Understanding and Use of Expert DNA Evidence," Australian Institute of Criminology Technical and Background Paper Series No. 37 (Canberra City: Australian Institute of Criminology, 2010), http://aic.gov.au/media_library/publications/tbp/tbp037/tbp037.pdf. The tutorials may be found at https://www.dropbox.com/s/piiedrwz3lqicfs/DNA_2008.mp4?dl=0 [DNA]; and https://www.dropbox.com/s/8kb4je9ia71eujn/RMP_2008.mp4?dl=0 [RMP].

51 Reality Technologies, "How Reality Technology is Used in Education," http://www.realitytechnologies.com/education.

52 Federal Judicial Center, "The Patent Process: An Overview for Jurors," video presentation, January 1, 2013, https://www.fjc.gov/publications/patent-process-overview-jurors.

53 Ibid.

54 Heather N. Mewe and Darren E. Donnelly, "Going to the Videotape: An Introduction to the Patent System" (Mountain View, Calif.: Fenwick and West, 2006), https://www.fenwick.com/FenwickDocuments/Going_to_Videotape.pdf.

55 John Gilleland, "The Debate is On: Is the Federal Judicial Center's Patent Tutorial Video Too Pro-Plaintiff?" *Illinois State Bar Association Newsletter*, March 2012, section on intellectual property law.

56 John D. Gilleland e-mail to Valerie P. Hans, June 6, 2017.

57 Shari Seidman Diamond, Mary R. Rose, and Beth Murphy, "Embedded Experts on Real Juries: A Delicate Balance," *William and Mary Law Review* 55 (2014): 885 – 933.

58 Diamond and Casper, "Blindfolding the Jury to Verdict Consequences" [see note 28].

59 Hans et al., "Science in the Jury Box," 68 [see note 39].

60 Richard O. Lempert, "Civil Juries and Complex Cases: Let's Not Rush to Judgment," *Michigan Law Review* 80 (1981): 68 – 132.

[61] Richard O. Lempert, "Civil Juries and Complex Cases: Taking Stock after Twelve Years," in *Verdict: Assessing the Civil Jury System*, ed. Robert E. Litan (Washington, D.C.: Brookings Institution Press, 1993), 181 – 247.

[62] Neil Vidmar and Valerie P. Hans, *American Juries: The Verdict* (Amherst, N.Y.: Prometheus Books, 2007), 68 – 69 [on the blue ribbon jury].

[63] Ibid., 69, 74 – 76. Nonetheless, jury pools still fall short of fully representing the community in many jurisdictions. See ibid., 76 – 81.

[64] Hans et al., "Science in the Jury Box" [see note 39]; David H. Kaye, Valerie P. Hans, B. Michael Dann, Erin Farley, and Stephanie Albertson, "Statistics in the Jury Box: How Jurors Respond to Mitochondrial DNA Match Probabilities," *Journal of Empirical Legal Studies* 4 (2007): 797 – 834; and B. Michael Dann, Valerie P. Hans, and David H. Kaye, "Can Jury Trial Innovations Improve Juror Understanding of DNA Evidence?" *National Institute of Justice Journal* 255 (2006): 2 – 6. See also Darrin R. Lehman, Richard O. Lempert, and Richard E. Nisbett, "The Effects of Graduate Training on Reasoning: Formal Discipline and Thinking about Everyday-Life Events," *American Psychologist* 43 (6) (1988): 431 – 442; Darrin R. Lehman and Richard E. Nisbett, "A Longitudinal Study of the Effects of Undergraduate Training on Reasoning," *Developmental Psychology* 26 (6) (1990): 952 – 960; Michael P. Weinstock, "Cognitive Bases for Effective Participation in Democratic Institutions: Argument Skill and Juror Reasoning," *Theory & Research in Social Education* 33 (1) (2005): 73 – 102; and Valerie F. Reyna and Charles J. Brainerd, "The Importance of Mathematics in Health and Human Judgment: Numeracy, Risk Communication and Medical Decision Making," *Learning & Individual Differences* 17 (2) (2007): 147 – 159.

[65] Ellen Peters, Daniel Vastfjall, Paul Slovic, et al., "Numeracy and Decision Making," *Psychological Science* 17 (5) (2006): 407 – 413.

[66] Diamond and Casper, "Blindfolding the Jury to Verdict Consequences" [see note 28].

[67] Valerie F. Reyna, Wendy L. Nelson, Paul K. Han, and Nathan F. Dieckmann, "How Numeracy Influences Risk Comprehension and Medical Decision Making," *Psychological Bulletin* 135 (6) (2009): 943 – 973.

[68] Kahneman, *Thinking, Fast and Slow* [general; see note 18]; Helm et al., "Trial by Numbers" [jurors; see note 32]; and Chris Guthrie, Jeffrey J. Rachlinski, and Andrew J. Wistrich, "Blinking on the Bench: How Judges Decide Cases," *Cornell Law Review* 93 (2007): 1 – 43 [judges].

[69] Rebecca Helm and James P. Dunlea, "Motivated Cognition and Juror Interpretation of Scientific Evidence: Applying Cultural Cognition to Interpretation of Forensic Testimony," *Penn State Law Review* 120 (1) (2016); and Dan Kahan, Ellen Peters, Maggie Wittlin, et al., "The Polarizing Impact of Scientific Literacy and Numeracy on Perceived Climate Change Risks," *Nature Climate Change* 2 (10) (2012): 732 – 735.

Bridging the Science-Law Divide

David Baltimore, David S. Tatel & Anne-Marie Mazza

Abstract: Formal opportunities for members of the scientific and legal communities to engage in ongoing collegial consideration of issues at the interface of science and law are limited. In the late 1990s, the National Academies of Sciences, Engineering, and Medicine established the Committee on Science, Technology, and Law (CSTL) – composed of equal numbers of members from science, engineering, and law – to provide an ongoing forum that would build permanent links between these communities. The range of issues investigated by the CSTL and the influence of these explorations are discussed in this essay.

DAVID BALTIMORE, a Fellow of the American Academy since 1974, is President Emeritus and the Robert Andrews Millikan Professor of Biology in the Division of Biology and Biological Engineering at the California Institute of Technology.

DAVID S. TATEL, a Fellow of the American Academy since 2015, is a Judge for the United States Court of Appeals for the District of Columbia Circuit.

ANNE-MARIE MAZZA is Senior Director of the Committee on Science, Technology, and Law at the National Academies of Sciences, Engineering, and Medicine.

(Complete author biographies appear at the end of the essay.)

Scientists and lawyers often appear to be speaking different languages. Each profession has its own culture and conventions, as well as its own jargon, and each employs distinctive means of resolving conflicts.[1] Often, when scientists and lawyers attempt to communicate, these differences can result in misunderstandings and confusion.[2] Moreover, when the institutions that represent these two professions attempt to collaborate, the likelihood of such difficulties can increase.

For almost two decades, the National Academies of Sciences, Engineering, and Medicine's Committee on Science, Technology, and Law (CSTL) has attempted to bridge the divide between the legal and scientific communities by developing projects and reports that encourage insightful consideration of scientific findings by legal institutions and appropriate oversight of the conduct of scientific, engineering, and biomedical research.[3] This essay discusses the origin of the CSTL and highlights some of the work the committee has undertaken to strengthen the bonds between science and law.[4]

The creation of a standing committee within the National Academies devoted to issues at the interface of science and law was not an easy decision. Many scientists within the National Academies viewed

doi:10.1162/DAED_a_00528

the sometimes brutal adversarial nature of the exchanges among legal professionals as unsuitable for an institution devoted to the scholarly search for scientific truth. The National Academies' mission of offering high-quality objective expert advice on some of the most pressing challenges facing the nation and the world seemed to some to be incompatible with the advocacy mission that animates much of legal discourse. When a need arose to address an issue pertinent to the legal profession, various committees of the National Academies would step up to offer advice on the particular situation, and then return to other issues focused more on scientific research than law.[5]

This ad hoc system of responding to issues involving questions of both science and law began to change in the 1990s. The science establishment could not help but recognize that science and law were becoming increasingly entangled in both the conduct of science and the development of public policy. Increasing regulation of scientific and academic research placed constraints on the conduct of scientific inquiry. Litigation was becoming more complex and often required testimony from scientific experts. Attorneys specializing in certain areas of litigation like toxic torts sometimes interpreted data, like clusters of adverse outcomes, in ways that were at odds with the consensus of the scientific community.

The legal system acknowledged the need for judges and attorneys to develop a more sophisticated understanding of science when, in the 1990s, the U.S. Supreme Court issued several decisions instructing judges to play a more active role in deciding what expert testimony a jury could hear. In the 1993 case *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, the Supreme Court stated that, in order to be admissible as evidence, scientific testimony must be based upon credible scientific methodology.[6] Judges were charged with conducting a rigorous assessment of the validity of scientific testimony before they decided to allow it. In establishing this standard, the Supreme Court quoted a brief submitted by the National Academy of Sciences and the American Association for the Advancement of Science as *amici curiae*: "Science is not an encyclopedic body of knowledge about the universe. Instead, it represents a *process* for proposing and refining theoretical explanations about the world that are subject to further testing and refinement."[7] The brief stated:

> The scientific community's well established criteria and institutional mechanisms for evaluating the validity of scientific assertions provide courts with clear and understandable guidance on how they can rationally and consistently evaluate scientific evidence. Courts should admit scientific evidence only if it conforms to scientific standards and is derived from methods that are generally accepted by the scientific community as valid and reliable. Such a test promotes sound judicial decision-making by providing workable means for screening and assessing the quality of scientific expert testimony in advance of trial.[8]

Several years later in 1999, in *Kumho Tire Co. v. Carmichael*, the Supreme Court considered whether "technical or other specialized knowledge," including testimony from the field of engineering, should also be evaluated for relevance and reliability in a manner consistent with the criteria offered in the *Daubert* decision. In that case, the National Academy of Engineering submitted an *amicus curiae* brief stating:

> Engineering, although differing in many respects from science, is founded on scientific understanding. In particular, the development of detailed understanding of the causes of the failure of an engineered device is a central feature of engineering: this effort involves a scientific-style investigation to understand the mechanism of failure at a fundamental, quantitative level.

In *Kumho*, the Supreme Court ruled that a trial court's gatekeeping role extends to all expert testimony.

In light of their important role in these Supreme Court decisions establishing the standards of admissibility for scientific evidence, the National Academies became more receptive to the establishment of an independent committee that could address topics that required an understanding of both science and law. The Academies' leadership came to agree with many leaders in the legal community (including U.S. Supreme Court Associate Justice Stephen Breyer and U.S. Federal District Court Judge Jack Weinstein) that there would be an ever-growing need for the legal and scientific communities to work with each other on issues of importance to the nation. The need for a prominent forum for representatives of these communities to get to know each other, understand their cultures, and exchange ideas was becoming more and more evident.

In March 2000, Donald Kennedy and Richard Merrill convened the Committee on Science, Technology, and Law, a new standing committee under the auspices of the National Academies of Sciences, Engineering, and Medicine.[9] Kennedy and Merrill sought to bring together distinguished members of the science and law communities to stimulate discussions that would lead to a better understanding of the role of science in legal decisions and government policies and to a better understanding of the legal and regulatory frameworks that govern the conduct of science. At biannual meetings, scientists and members of the legal community, including members of the legal academy and judiciary, were encouraged to bring to the committee topics of national importance that would be best addressed from the perspective of both communities. Sessions at each meeting were built around controversial or emerging issues and often led to the development of project ideas for consensus studies and convening activities. At the time it was established, Kennedy and Merrill noted, the CSTL could "not hope to canvass the entire terrain. Instead, we hope to become one of several contributors to the growing dialogue between science, engineering, and law; a supporter of initiatives by other organizations; and a catalyst for promoting productive collaboration among participants from all affected disciplines."[10] Eighteen years later, it's probably fair to say that Kennedy and Merrill could never have envisaged either the wide range of topics the CSTL would explore or the impact of these explorations.

In 2009, Kennedy and Merrill passed leadership of the CSTL to Richard Meserve and David Korn, and in 2015, Meserve and Korn passed leadership of the committee to David Baltimore and David S. Tatel (coauthors of this essay).[11] It is clear that the National Academies' and Kennedy and Merrill's decision to establish the CSTL was prescient. Many issues we face today, and will face in the future, require and benefit from the active engagement of both the legal and scientific communities.

The pursuit of truth is a goal of both science and law. Science is almost always open-ended: it is a process for investigating nature that reaches tentative interpretations based on the data at hand and subject to reinterpretation as continuing investigations generate ever more data that modify prior understandings or provide new insight. The law, too, requires constant refinement and reinterpretation. From both professions, society often demands that practical decisions be made on the basis of incomplete information. Both scientists and judges seek explanations for phenomena to gain a better understanding of a particular situation. The scientist seeks truth through an iterative process wherein a hypothesis is posed, data are collected and analyzed, and new understanding is gained that then gener-

ates new lines of inquiry. While this may appear to be a straightforward linear process, in practice, science is often surprisingly messy. Concurrent parallel lines of inquiry, collaborative exchange, and ongoing efforts to build a consensus through review and commentary on emerging research are the norm. The result in the best-case scenario is a transparent process that provides an opportunity for correction and refinement through peer review and further study. In science, evidence is continuously gathered, challenged, and refined until consensus develops, though a degree of uncertainty is associated with most scientific conclusions. The scientific community readily accepts that today's knowledge could be (and should be) revised if new data and findings lead to new conclusions. By continuing to collect evidence and test the limits of theories, the scientific enterprise, by its very nature, is self-correcting.

Law also builds on the past, though change proceeds at a more deliberate pace. In interpreting the law and in some admissibility decisions, precedent is given great weight, and judges typically do not have the option of postponing judgment until additional information emerges. In areas ranging from climate and the environment to medical practice and pharmacology, regulations and laws are written even though scientific understanding may be incomplete and uncertain. Legal disputes must be resolved without delay based upon the data at hand, and the legal community must respect legal conventions that may constrain the search for truth. As noted in *Ethyl Corporation v. Environmental Protection Agency*,

> We must look at the decision not as the chemist, biologist or statistician that we are qualified neither by training nor experience to be, but as a reviewing court exercising our narrowly defined duty of holding agencies to certain minimal standards of rationality. Although (our) inquiry into the facts is to be searching and careful, the ultimate standard of review is a narrow one. We must affirm unless the agency decision is arbitrary or capricious.[12]

While relying on legal precedent established by previous rulings may provide continuity, such precedent may impede consideration of advances in science and technology as they emerge outside the courtroom.

The difficulty in reforming common legal practice became apparent when, in 2006, under the CSTL's auspices, a committee was appointed to undertake a congressionally mandated study of the forensic sciences in the United States. The committee – which was cochaired by Judge Harry Edwards, U.S. Court of Appeals for the District of Columbia Circuit, and Constantine Gatsonis, professor and director of the Center for Statistical Sciences at Brown University – issued a landmark study in 2009 that found the forensic sciences to be systemically flawed. In reviewing the scientific underpinnings of many forensic disciplines, the committee found a lack of rigorous scientific research and noted that "[w]ith the exception of nuclear DNA analysis . . . no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source."[13] This finding undercuts decades of testimony by forensic experts who asserted that evidence associated with a crime scene could have originated with only one individual or object, to the exclusion of all other persons or objects in the world.[14]

For example, one of the forensic techniques that the committee reviewed in detail was forensic odontology, or "bite mark" analysis. Like many of the forensic sciences, the criteria for evaluating bite mark evidence were developed in the context of law enforcement investigations and not in scientific laboratories. In 1974, three dentists testified that they could match bite impres-

sions found on the body of an exhumed victim to a suspect, Walter Edgar Marx. Marx was convicted, and the decision was upheld by a California appeals court in 1975. Despite the unprecedented nature of the claims, the California appeals court declined to require a *Frye* hearing, since the techniques used by the dentists (for instance, X-rays, models, microscopy, and photography) were well established.[15] Since there was no scientific methodology or testing of bite marks, nor any science connecting an individual to bite marks, the court concluded there was therefore no need for such a hearing. Instead, the appeals court deferred to the trial judge who believed that bite mark evidence was sound. The Marx decision laid the foundation for the admission of bite marks into evidence and set a precedent that has influenced many courts, despite the fact that bite mark evidence has now largely been discredited.

In conducting its review, the committee found that

> Although the majority of forensic odontologists are satisfied that bite marks can demonstrate sufficient detail for positive identification, no scientific studies support this assessment, and no large population studies have been conducted. In numerous instances, experts diverge widely in their evaluations of the same bite mark evidence, which has led to questioning the value and scientific objectivity of such evidence.

Further, the committee noted that it "received no evidence of an existing scientific basis for identifying an individual to the exclusion of all others." Following such an assessment, one might expect that judges would no longer allow testimony that links a bite mark to a specific individual. Yet, today, bite mark evidence remains admissible in some courts,[16] although not in others.[17] The scientific community finds the resistance to change by the legal community difficult to understand. Scientists are mystified when the courts resist embracing new knowledge that represents a scientific consensus. It appears that the law has few systems designed to take advantage of the evolving nature of scientific knowledge. Judges seem to feel free to ignore scientific advances, especially in cases involving forensic sciences.[18] Procedures to aid scientific understanding in civil trials, such as *Daubert* hearings, appear to be far rarer in criminal trials, although the Court in *Daubert* was interpreting a federal rule of evidence that is in theory equally applicable in civil and criminal litigation. Research is needed to understand the extent to which *Daubert* hearings are held in civil and criminal cases. If there is a wide variation, it would be important to understand whether this is a failure of judicial education, or a reflection of a more systemic issue.[19]

Cases involving complex scientific evidence can place great demands on judges. Most judges do not come to the bench with a strong background in science or technology. Following the *Daubert* decision, judges needed more information on how scientists determine the validity of scientific assertions. In an effort to provide such assistance, the Federal Judicial Center (FJC), the research and education agency of the federal judiciary, developed a series of educational programs that allowed judges to interact with scientists to better understand the culture, process, and methods of science. In 1995, the FJC developed the first edition of the *Reference Manual on Scientific Evidence*, in part in response to the Supreme Court's *Daubert* decision. In order to satisfy *Daubert*'s reliability standards, the Supreme Court instructed judges to consider whether a proffered expert opinion was the product of scientific reasoning and scientifically sound methodology. The chapters in the *Reference Manual* describe basic principles of major scientific fields from which legal evidence is typically derived and provide examples of cases where such evidence was

used. The manual contains glossaries of technical terms that scientists may use in particular areas of scientific inquiry.

The purpose of the *Reference Manual* is to provide judges with sufficient understanding to hold an informed conversation with expert witnesses and attorneys while considering challenges to the admissibility of the scientific evidence. As noted in *Lead Industries Association, Inc. v. Environmental Protection Agency*:

> [T]he court "must understand enough about the problem confronting the agency to comprehend the meaning of the evidence relied upon and the evidence discarded; the questions addressed by the agency and those bypassed; the choices open to the agency and those made." . . . However, it is appropriate to sound some notes of caution about the limits of this exercise. First, we would be less than candid if we failed to acknowledge that we approach the task of examining some of the complex scientific issues presented in cases of this sort with some diffidence. More important, we stress that our review of the evidence is not designed to enable us to second-guess the Agency's expert decisionmaker. . . . Congress has entrusted the Agency with the responsibility for making these scientific and other judgments, and we must respect both Congress' decision and the Agency's ability to rely on the expertise that it develops.[20]

Soon after the establishment of the CSTL, the National Academies and the FJC recognized a unique opportunity to establish stronger ties between the scientific community and the federal judiciary. The director of the FJC and the program officer overseeing the Center's studies on scientific evidence were given permanent memberships on the committee. The FJC subsequently asked the CSTL to collaborate on the development of an expanded third edition of the *Reference Manual on Scientific Evidence*.

An advisory committee – cochaired by Judge Gladys Kessler, U.S. District Court for the District of Columbia, and Jerome Kassirer, professor of medicine at Tufts University, and including judges, scientists, engineers, and medical professionals – was charged with overseeing the development of reference guides on thirteen scientific, engineering, and medical topics. As in the previous editions, the guides were designed to assist judges as they attempted to assess the scientific foundation of scientific testimony and, in the absence of a jury, to adjudicate on differing interpretations of scientific evidence. In addition to updating previous guides, the expanded version included new chapters on topics such as neuroscience, forensics identification, exposure science, and mental health. The *Reference Manual* is provided to more than three thousand federal judges and is also widely used by state judges, attorneys, and law professors. The National Academies makes the *Reference Manual* available for free to the public, and it has become one of the forty most-downloaded reports of the over 9,900 reports issued by the National Academies, with 30 percent of the downloads coming from nations other than the United States. As Justice Breyer noted in the introduction to the volume, "This manual seeks to open legal institutional channels through which science – its learning, tools, and principles – may flow more easily and thereby better inform the law. The manual represents one part of a joint scientific-legal effort that will further the interests of truth and justice alike."[21]

Academic research in the United States is governed by a host of laws, regulations, and policies that provide oversight of scientists and engineers who conduct research using taxpayer dollars. The CSTL has evaluated numerous regulations and policies affecting scientific research and research institutions. In its early years, the CSTL became interested in government policies affecting access to, and the use and evalua-

tion of, research findings or scientific findings relied upon by government agencies. Questions arose as to the applicability of *Daubert* to administrative agencies.[22] This interest was stimulated by the Data Quality Act, which directed the White House Office of Management and Budget (OMB) to develop government-wide guidelines to "provide policy and procedural guidance to federal agencies for ensuring and maximizing the quality, objectivity, utility, and integrity of information . . . disseminated by Federal agencies."[23] Companion guidelines included in the OMB Information Quality Bulletin for Peer Review provided guidance to agencies regarding how to conduct peer review of the "most important science disseminated by the Federal Government."[24] Both the research community and federal agencies expressed concerns, however, that the proposed OMB guidelines would, in the name of quality and transparency, disrupt scientific practice and would be used by special interest groups to contest the scientific premises of government rules and regulations. To some, the Data Quality Act appeared to be a *Daubert*-like screening of scientific information and agency processes relied upon by federal regulatory agencies.[25] At the request of the OMB, in 2002 and 2003, the CSTL convened a series of workshops where the affected communities (federal agencies, researchers, public interest groups, and industry) could express their concerns to the OMB. John Graham, the administrator of the OMB's Office of Information and Regulatory Affairs (OIRA), and several other OMB senior staff attended these workshops. The exchange of information during these discussions led to substantive revisions to both sets of OMB guidelines and an apparently greater understanding of the scientific process by the OIRA.

Most recently, in 2016, the CSTL issued a report entitled *Optimizing the Nation's Investment in Academic Research: A New Regulatory Framework for the 21st Century*. The report was commissioned by Congress and authored by a CSTL study committee chaired by Larry Faulkner, president emeritus of the University of Texas at Austin; Harriet Rabb, general counsel of The Rockefeller University, served as vice chair. The report considers a broad range of regulations governing academic research, from proposal development to the acceptance of an award, to the conduct of research, to the final closeout of a contract or grant. The study recognizes the importance of regulation in protecting the government, research institutions, investigators, and the public from fraud, waste, and abuse, while providing an organizing framework for the conduct of research. The report found that the increasing number of laws, regulations, and policies emerging over past decades have had the unintended negative effect of diverting significant researcher time from research. In essence, the country is not reaping the full benefits from all the research it is funding:

> The continuing expansion of federal regulations and requirements is diminishing the effectiveness of the U.S. research enterprise and lowering the return on the federal investment in basic and applied research by diverting investigators' time and institutional resources away from research and toward administrative and compliance matters. A new framework . . . is needed to ensure that regulatory requirements are justified, proportional to the problems being addressed, and harmonized across funding agencies so as to create a more effective and efficient partnership between funding agencies and research institutions.[26]

Among its many recommendations, the report called for the establishment of a Research Policy Board as an "analytical, anticipatory, and coordinating forum on research regulatory policy." This recommendation, along with many of the committee's other recommendations, was enacted with the passage of several laws, including the 21st Century Cures Act (2016).[27]

David Baltimore, David S. Tatel & Anne-Marie Mazza

The committee also called upon Congress to appoint a national commission on human research subjects and recommended that the Department of Health and Human Services withdraw its Notice of Proposed Rulemaking on the Federal Policy for the Protection of Human Subjects. The committee argued that since the 1978 issuance of the Belmont Report, which articulated three principles key to the protection of human beings used in research studies,

> the biomedical and sociobehavioral research enterprises have grown enormously. This growth, accompanied by the development of a remarkable number of new research capabilities and contexts, raises questions as to the optimum application and balancing of the Belmont principles, as well as whether these principles are, in and of themselves, still sufficient pillars upon which to build human research protection programs and regulations. In addition, the overarching legal and regulatory frameworks and institutional arrangements governing human research subjects require reconsideration and clarification.[28]

While a commission has not been appointed and the final rule was issued in January 2017, it is important to note that newly issued federal policy reflects many of the committee's concerns.[29] Most notably, it did not adopt a controversial proposal to require researchers to obtain informed consent to use unidentified biospecimens in research.

CSTL activities demonstrate the importance and value of having the legal and scientific communities involved in the development of the legal and regulatory apparatuses that govern research and in discussions about how scientific research is conducted. A better understanding of both cultures affords the nation an opportunity to maximize the value of its considerable investment in research for the benefit of the American economy and the health and social well-being of its citizens.

Regulation of emerging technologies has been of particular interest to CSTL members. The committee has convened meetings on synthetic biology, gain of function research, neuroscience, and human genome editing, to name just a few topics. In the course of these explorations, it has become increasingly clear that law and science speak only to some of the issues that arise, and that it is imperative to consider ethical frameworks as well. As emerging technologies become a more prominent part of public discussions, ethical, moral, and societal issues must be part of future public dialogues.

We have also learned that it is necessary to expand our discussions beyond just the United States to include colleagues from around the world. At the behest of the National Academy of Sciences (NAS) and National Academy of Medicine (NAM), the CSTL led the Academies' collaboration with the Royal Society and Chinese Academy of Sciences to organize an international summit on human genome editing.[30] A new gene editing tool, CRISPR-Cas9, captured the public's attention in 2015 when it became known that this tool could be used to alter the human germline. The use of CRISPR-Cas9 to edit human genes raises profound questions about the manner in which the DNA of living beings may be altered, as well as the genomes of future offspring. The two-and-a-half day summit, held in December 2015, received worldwide attention, with representatives from more than twenty countries in attendance. The live webcast attracted more than three thousand viewers from seventy-one nations. At the conclusion of the summit, the summit planning committee released a statement, "On Human Gene Editing":

> It would be irresponsible to proceed with any clinical use of germline editing unless and until (i) the relevant safety and efficacy issues have been resolved, based on appropriate understanding and balancing of risks, poten-

tial benefits, and alternatives, and (ii) there is broad societal consensus about the appropriateness of the proposed application. Moreover, any clinical use should proceed only under appropriate regulatory oversight.[31]

Recognizing that the human genome is "shared among all nations," the statement called for "an ongoing international forum to discuss potential clinical uses of gene editing; help inform decisions by national policymakers and others; formulate recommendations and guidelines; and promote coordination among nations."[32] Since the 2015 summit, the NAS and NAM issued a consensus report, *Human Genome Editing*, that indicated that, in the future, clinical trials for genome editing of the human germline could be permitted, but only for serious conditions under stringent oversight. The report outlines several criteria that should be met before allowing such trials to go forward.[33] Other organizations have issued guidance as well.[34] A second international summit co-organized by the NAS, NAM, the Royal Society, and the Academy of Sciences of Hong Kong will be held in Hong Kong in November 2018.[35]

As the CSTL continues to chart its course, it has identified several important topics where science, law, ethics, and international engagement will play a critical role.

In recent decades, major advances in neuroscience, psychology, behavioral economics, and related fields have expanded our understanding of human cognition and mental processes. This work has had wide-ranging significance in illuminating phenomena such as visual perception, memory, rational choice, and decision-making. A 2014 CSTL report, *Identifying the Culprit: Assessing Eyewitness Identification*, chaired by Thomas D. Albright, director of the Vision Center and Laboratory at the Salk Institute for Biological Studies, and Judge Jed Rakoff, U.S. District Court for the Southern District of New York, synthesized and applied insights from this body of research to the topic of eyewitness identifications. The report discussed the scientific foundations of visual perceptual experience and memory, identified key factors that can lead to error, and offered recommendations for best practices to improve the accuracy of eyewitness identification in criminal investigations. In 2017, the U.S. Department of Justice issued new procedures for how the FBI and other law enforcement agencies should ask eyewitnesses to identify suspects using photo lineups.[36]

The related topic of unconscious bias has garnered much attention in light of well-publicized incidents of police use of force against minority citizens. While this is hardly the first time the issue has been at the forefront of national conversation, today we can engage in this conversation against a backdrop of over two decades of scientific research on the cognitive mechanisms that underlie unconscious bias. The CSTL envisions a study that recognizes the pervasiveness of unconscious bias as a common aspect of mental processing in a wide variety of contexts affecting a wide variety of groups. For example, recent studies by social psychologist Kelly M. Hoffman and colleagues have demonstrated that medical students and residents who held false beliefs regarding biological differences between blacks and whites (for example, that black people's skin is thicker than white people's skin) showed racial bias in the accuracy of not only their pain assessments, but also their treatment recommendations.[37] In a separate context, then–acting director of the U.S. Office of Personnel Management, in congressional testimony offered in 2016, described gender bias in federal hiring practices and identified unconscious bias as the most challenging barrier to diversity and inclusion.[38]

At a recent speech to incoming students at Georgetown Law, Associate Supreme

Court Justice Ruth Bader Ginsberg observed that, while many overt barriers to employment discrimination are gone, what remains is unconscious bias, which is harder to address. Unconscious biases have profound implications for efforts to increase diversity in hiring and promotion practices across all sectors of the economy, for criminal justice, and for decisions regarding housing and finance.

Scientific understanding of unconscious bias has advanced considerably in recent decades, but this body of research has had minimal impact on law and policy.[39] Some courts and judges have occasionally recognized the reality of unconscious bias. As Justice Kennedy noted: "Recognition of disparate-impact liability under the [Fair Housing Act] also plays a role in uncovering discriminatory intent: It permits plaintiffs to counteract unconscious prejudices and disguised animus that escape easy classification as disparate treatment."[40] Nonetheless, there is no systematic or well-developed approach to how such biases should be taken into account under relevant legal standards.

The emerging body of research on unconscious bias has the potential to inform and motivate institutional reform in multiple environments. In police departments, universities, industry, and other settings, administrators are searching for ways to reduce unconscious bias not only to lessen legal exposure, but also to achieve diversity-related objectives and to improve organizational performance and credibility. By bringing the disciplined focus of science to bear on this critical issue, the CSTL seeks to stimulate new conversations about the nature of discrimination and to identify ways to counteract ingrained unconscious modes of information processing.

To take another example, the exploration of outer space has until recently been the exclusive domain of a few prominent governments. Today, however, we are witnessing increasing interest in the exploration of space by emerging nations and nongovernmental entities. Scientific and technological advances – such as the development of small satellites for research, communications, and remote sensing, and commercial launch services – are rapidly changing access to space and expanding the scope of space activities. The diversification and growth of new actors and activities in space raise questions about the adequacy of existing laws, regulations, and policies. As recently noted by Joan Johnson-Freese,

> Fifty years on, the Outer Space Treaty and its spin-offs are still appropriate. But interpretations of its provisions are, more than ever, being influenced by commercial interests and politics. Supplementary rules and norms are needed. In an era in which international cooperation on treaties is tenuous, informal agreements and resolutions must guide space-faring actors, protect the environment and prevent wars.[41]

The CSTL sees this new era of activity in space as an appropriate time to explore and evaluate the adequacy of the legal, policy, and regulatory regimes governing the exploration and use of space.

In this essay, we have taken the opportunity to describe the history of the CSTL, to provide some examples of the work the committee has done, and to identify areas of concern that will be the topics of study in the future. We have tried to illustrate the richness that emerges from thinking about the interface of science, technology, and law. Interestingly, both science and law have the same property of never being fixed and complete. It is our hope that having members from the worlds of science and law meet regularly provides a venue in which viewpoints are broadened on a range of issues, thus furthering understanding in both communities that extends beyond individual committee members.

As the knowledge of science and the procedures of law evolve, the need for this "cross-pollination" becomes ever more necessary. Thus, the work of the CSTL will never be complete, and its particular concerns will inevitably vary over time.

AUTHOR BIOGRAPHIES

DAVID BALTIMORE, a Fellow of the American Academy since 1974, is President Emeritus and the Robert Andrews Millikan Professor of Biology in the Division of Biology and Biological Engineering at the California Institute of Technology. He is Cochair of the Committee on Science, Technology, and Law at the National Academies of Sciences, Engineering, and Medicine. His recent publications include articles in the journals *Blood*, *Journal of Immunotherapy*, and *Proceedings of the National Academy of Sciences*.

DAVID S. TATEL, a Fellow of the American Academy since 2015, is a Judge for the United States Court of Appeals for the District of Columbia Circuit. He is Cochair of the Committee on Science, Technology, and Law at the National Academies of Sciences, Engineering, and Medicine.

ANNE-MARIE MAZZA is Senior Director of the Committee on Science, Technology, and Law at the National Academies of Sciences, Engineering, and Medicine. The committee's recent publications include *Dual Use Research of Concern in the Life Sciences: Current Issues and Controversies* (2017), *Making the Living World Engineerable: Science, Practice, and Policy* (2016), *Optimizing the Nation's Investment in Academic Research* (2016), and *Strengthening Forensic Science in the United States: A Path Forward* (2009).

ENDNOTES

[1] "[L]aw and science are both knowledge-generating institutions, but . . . fact-making serves different functions in these two settings." Sheila Jasanoff, "Law's Knowledge: Science for Justice in Legal Settings," *American Journal of Public Health* 95 (S1) (2005): S49 – 59.

[2] As Susan Haack put it: "two simple observations: that the work of a scientist is very different from the work of an attorney or a judge; and that, when scientists give expert testimony or advise a court, a regulative body, etc., communication can be difficult and very imperfect." Susan Haack, "Scientific Inference vs. Legal Reasoning? – Not so Fast!" paper shared at the *Dædalus* authors' conference on "Science and the Legal System," July 2017, Cambridge, Massachusetts.

[3] The CSTL is not the first attempt at building better understanding and communication between these two communities. Indeed, Sheila Jasanoff was, and remains, a pioneer in this field, providing foundational understanding of the interactions between law and science. Also, in 1974, the American Association for the Advancement of Science and American Bar Association founded the National Conference of Lawyers and Scientists. See American Association for the Advancement of Science, "National Conference of Lawyers and Scientists," https://www.aaas.org/page/national-conference-lawyers-and-scientists. And, in 1975, the National Research Council established the Committee on Law and Justice "to improve governmental decision making and public policy, and promote the understanding and dissemination of research in matters involving law and justice." See The National Academies of Science, Engineering, and Medicine, http://sites.nationalacademies.org/DBASSE/CLAJ/DBASSE_073357.

[4] Throughout this essay, when we refer to the scientific and legal communities, we mean to include with the scientific community the engineering and medical communities, and when we refer to the legal community, we are including the judiciary, legal academy, legal practitioners, and policy-makers. (For a complete view of the CSTL's work, see www.nationalacademies.org/stl.)

[5] See, for example, the National Academies' reports: National Research Council, *DNA Technology in Forensic Sciences* (Washington, D.C.: National Academies Press, 2004); and National Re-

search Council, *Forensic Analysis: Weighing Bullet Lead Evidence* (Washington, D.C.: National Academies Press, 2004).

[6] *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993). In *Daubert*, the Supreme Court ruled that a "trial judge must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable." To assist in this assessment, the Court offered some guideposts for identifying proffered scientific testimony that was properly grounded in scientific knowledge. The Court suggested that trial judges ought to consider: 1) whether a theory or technique "can be (and has been) tested"; 2) whether the theory or technique "has been subjected to peer review and publication"; 3) "the known or potential rate of error" of a particular scientific technique; 4) "the existence and maintenance of standards controlling the technique's operation"; and 5) a scientific technique's degree of acceptance within a relevant scientific community. The Supreme Court further developed these standards in *General Electric Co. v. Joiner*, 522 U. S. 136 (1997); and *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999). It should be pointed out, however, that the Court was careful to make clear that meeting these suggested standards was neither necessary nor necessarily sufficient to justify the admission of scientific evidence. Nonetheless, lower courts have often acted as though they were.

[7] Brief for the American Association for the Advancement of Science and the National Academy of Sciences as *amici curia* in Support of Respondent in *Daubert v. Merrell Dow Pharmaceuticals Inc.*, January 19, 1993. Emphasis in original.

[8] Ibid., 27. The brief specified criteria such as the generation and testing of hypotheses, replication of results, the use of peer review, and acceptance by the scientific community.

[9] Donald Kennedy is president emeritus of Stanford University, administrator emeritus of the U.S. Food and Drug Administration, and emeritus editor-in-chief of *Science* magazine. Richard Merrill is formerly of counsel for Covington and Burling, general counsel for the U.S. Food and Drug Administration, and dean emeritus of the School of Law at the University of Virginia.

[10] Donald Kennedy and Richard A. Merrill, "Science and the Law," *Issues in Science and Technology* 16 (4) (2000): 49 – 51.

[11] Richard Meserve is senior of counsel for Covington and Burling and president emeritus of the Carnegie Institution for Science. It should also be noted that he coauthored *amici curia* briefs submitted on behalf of the National Academies in both *Daubert v. Merrell Dow Pharmaceuticals* and *Kumho Tire Co. v. Carmichael*. David Korn is professor of pathology at Massachusetts General Hospital and Harvard Medical School, and dean emeritus at the Stanford University Medical School.

[12] *Ethyl Corporation v. Environmental Protection Agency*, 541 F.2d 1, 36 – 37 (D.C. Cir. 1976).

[13] National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (Washington, D.C.: National Academies Press, 2009), 7.

[14] The report stimulated an ongoing national discussion on the need to improve forensic science; prompted the creation of a DOJ–NIST National Commission on Forensic Science and the establishment by NIST of a Forensic Science Center of Excellence; led the FBI to review thousands of cases where testimony regarding hair evidence was suspect; generated multiple proposals of federal legislation adopting the report's recommendations; has been cited in numerous court decisions, including decisions by the U.S. Supreme Court (see Antonin Scalia in *Melendez-Diaz v. Massachusetts*: "According to the CSTL forensic science report, '[t]he majority of [laboratories producing forensic evidence] are administered by law enforcement agencies, such as police departments, where the laboratory administrator reports to the head of the agency.' And '[b]ecause forensic scientists often are driven in their work by a need to answer a particular question related to the issues of a particular case, they sometimes face pressure to sacrifice appropriate methodology for the sake of expediency.' A forensic analyst responding to a request from a law enforcement official may feel pressure – or have an incentive – to alter the evidence in a manner favorable to the prosecution."); and has generated ongoing media and popular cultural coverage (see, for example, *Last Week Tonight with John Oliver*, October 1, 2017, https://www.youtube.com/watch?v=ScmJvmzDcGo). Nonetheless, progress is slow and efforts to thwart progress continue: in 2017, the DOJ, under the direc-

tive of Attorney General Jeff Sessions, decided not to renew the National Commission, opting instead to appoint a former prosecutor as the DOJ's forensic science advisor.

[15] In *Frye v. United States*, the U.S. Court of Appeals for the D.C. Circuit ruled that to be admitted into court, scientific evidence must have "gained general acceptance in the particular field in which it belongs." *People v. Marx*, 54 Cal. App. 3d 100 (2d Dist. 1975). For a more detailed discussion of *People v. Marx*, see David L. Faigman, Edward K. Cheng, Jennifer L. Mnookin, et al., *Modern Scientific Evidence: The Law and Science of Expert Testimony*, 2016 – 2017 ed. (Eagan, Minn.: Thomson Reuters, 2016), sec. 35:5.

[16] See, for example, Honorable Jolene Grubb Kopriva, March 8, 2017, Opinion and Order in *Commonwealth of Pennsylvania v. Paul Allen Ross*, in which Judge Kopriva, in refusing to exclude bite mark evidence, wrote, "Although the use of bite mark evidence is beginning to face challenges, it would be premature for this court to order that the methodology is no longer generally accepted in the relevant scientific community."

[17] In 2016, the Texas Forensic Science Commission recommended a moratorium on the use of bite mark evidence in future criminal prosecutions in Texas until the technique could be scientifically validated, and ordered a review of every conviction in Texas in which bite marks were used.

[18] See Stephanie Damon-Moore, "Trial Judges and the Forensic Sciences Problem," *NYU Law Review* 92 (2017): 1570; "Trial judges are uniquely well positioned to staunch the flow of unreliable forensic evidence into court. . . . In order to do so, however, trial judges must break with sometimes-lengthy histories of admission, engage in a technical analysis outside the wheelhouse of most lawyers, and perhaps even face political backlash against an unpopular decision. As difficult as this may seem, none of the obstacles facing trial judges are insurmountable, and none exempt trial judges from their obligation to vigilantly gatekeep expert evidence in their courtrooms. . . . Now more than ever, trial judges must lead the way toward a better future for forensic evidence."

[19] See Peter J. Neufeld, "The (Near) Irrelevance of *Daubert* to Criminal Justice and Some Suggestions for Reform," *American Journal of Public Health* 95 (S1) (2005): S107, S109.

[20] *Lead Industries Association, Inc. v. Environmental Protection Agency*, 647 F.2d 1130, 1145 – 1146 (D.C. Cir. 1980).

[21] National Academy of Sciences, *Reference Manual on Scientific Evidence*, 3rd ed. (Washington, D.C.: National Academies Press, 2001), 9.

[22] Wendy E. Wagner, "Importing *Daubert* to Administrative Agencies through the Information Quality Act," *Journal of Law and Policy* 12 (2) (2004).

[23] Section 515(a) of the Treasury and General Government Appropriations Act for Fiscal Year 2001 (Public Law 106-554).

[24] Office of Management and Budget, Executive Office of the President, "Final Information Quality Bulletin for Peer Review," *Federal Register* 70 (10) (January 14, 2005).

[25] Wagner, "Importing *Daubert*," 590 – 591 [see note 22].

[26] National Academies of Sciences, Engineering, and Medicine, *Optimizing the Nation's Investment in Academic Research* (Washington, D.C.: National Academies Press, 2016), 6.

[27] 21st Century Cures Act (Public Law 114-255).

[28] National Academies of Sciences, Engineering, and Medicine, *Optimizing the Nation's Investment in Academic Research*, 167 [see note 26].

[29] Its implementation has been twice postponed, and its full implementation is now scheduled for January 21, 2019. See Jerry Menikoff, Julie Kaneshiro, and Ivor Pritchard, "The Common Rule, Updated," *New England Journal of Medicine* 376 (2017): 613 – 615, wherein the authors state: "influential reports, including one from the National Academies of Sciences, Engineering, and Medicine, led to a long process of deliberation and discussion. The result is a final rule that differs significantly from what was initially proposed."

[30] National Academies of Science, Engineering, and Medicine, "International Summit on Human Gene Editing," http://nationalacademies.org/gene-editing/Gene-Edit-Summit/.

[31] Organizing Committee for the International Summit on Human Gene Editing, "On Human Gene Editing: International Summit Statement," December 3, 2015, http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=12032015a.

[32] Ibid.

[33] National Academies of Sciences, Engineering, and Medicine, *Human Genome Editing: Science, Ethics, and Governance* (Washington, D.C.: The National Academies Press, 2017).

[34] See Nuffield Council on Bioethics, *Genome Editing and Human Reproduction: Social and Ethical Issues* (London: Nuffield Council on Bioethics, 2018), http://nuffieldbioethics.org/wp-content/uploads/Genome-editing-and-human-reproduction-FINAL-website.pdf; and Association for Responsible Research and Innovation in Genome Editing (ARRIGE), https://arrige.org/.

[35] See The National Academies of Science, Engineering, and Medicine, "Second International Summit on Human Genome Editing," http://nationalacademies.org/gene-editing/2nd_summit/index.htm.

[36] United States Department of Justice, "Eyewitness Identification: Procedures for Conducting Photo Arrays," January 6, 2017, https://www.justice.gov/file/923201/download.

[37] Kelly M. Hoffman, Sophie Trawalter, Jordan R. Axt, and M. Norman Oliver, "Racial Bias in Pain Assessment and Treatment Recommendations and False Beliefs about Biological Differences Between Blacks and Whites," *Proceedings of the National Academy of Sciences* 113 (16) (2016): 4296 – 4301.

[38] Joe Davidson, "Feds Urged to Fight 'Unconscious Bias' in Hiring and Promotions," *The Washington Post*, April 14, 2016, https://www.washingtonpost.com/news/powerpost/wp/2016/04/14/feds-urged-to-fight-unconscious-bias-in-hiring-and-promotions/.

[39] See generally Jerry Kang, Mark Bennett, and Devon Carbado, "Implicit Bias in the Courtroom," *UCLA Law Review* 59 (2012): 1124; and Jerry Kang and Kristin Lane, "Seeing through Colorblindness: Implicit Bias and the Law," *UCLA Law Review* 58 (2010): 465.

[40] *Texas Department of Housing and Community Affairs v. Inclusive Communities Project*, 576 U.S. ___ (2015).

[41] Joan Johnson-Freese, "Build on the Outer Space Treaty," *Nature* 550 (7675) (2017): 182 – 184.

Inside back cover: Building bridges between science and the law is an ongoing activity. Progress is not always fast, but it is real. Courts increasingly attend to science, and some members of both the legal and science communities are committed to working together to improve the quality and comprehensibility of the science that courts receive. We are optimistic.


100ml
80
60
40

on the horizon:

Why Jazz Still Matters
edited by Gerald Early & Ingrid Monson

with Krin Gabbard, Farah Griffin, Vijay Iyer, Kelsey Klotz, Carol Muller, Robert O'Meally, Gabriel Solis, Jason Stanyek, Judith Tick & Christopher J. Wells

Access to Justice
edited by Lincoln Caplan, Lance Liebman & Rebecca L. Sandefur

Inequality as a Multidimensional Process
edited by Michèle Lamont & Paul Pierson

Representing the intellectual community in its breadth and diversity, **Dædalus** *explores the frontiers of knowledge and issues of public importance.*

U.S. $15; www.amacad.org;
@americanacad