A Preface to the Seventy-fifth Anniversary Volume

This Volume marks a milestone in the history of the Brooklyn Law Review—its seventy-fifth year of publication. The Law Review was actually founded in the fall of 1931—more than seventy-eight years ago—and its first issue came out in the spring of 1932. The mismatch in the years is explained by the fact that publication was suspended during part of World War II.

An event like this is an opportunity to reflect on our past and to thank those who have contributed to the Law Review’s success and growth over the years. When the Law Review was first taking shape during the 1931-32 academic year, the country was in the midst of the Great Depression. Brooklyn Law School was still part of St. Lawrence University; the school’s founder, William Payson Richardson, was Dean; and future Dean Jerome Prince was on the journal’s original staff, soon to become its first student Editor-in-Chief. The masthead counted only nineteen editors, and the inaugural issue featured a portrait of Benjamin N. Cardozo to celebrate the famous New York judge’s recent elevation to the United States Supreme Court.

Today’s Law Review is a product of a different era. We have recently witnessed the election of this country’s first African-American president, the unfolding of what has been described as the worst economic crisis in seven decades, and the appointment of two more New Yorkers to the Supreme Court. Our membership has grown to eighty-five students. We count among our alumni federal and state judges, accomplished practitioners, and, to the delight of our members, several past and present Brooklyn Law School professors. We have kept our generalist scope but almost every volume these days includes a special symposium issue, which, according to Dean Prince’s
Preface to the twenty-fifth anniversary volume, was something the Law Review was just beginning to experiment with in the 1950s. More importantly, however, just like in our early years, today, we continue to serve as an important national forum for intellectual debate, a driver for legal reform, and an incubator for student leadership and camaraderie.

On behalf of all my fellow members, I’d like to thank all of the past editors, our faculty advisors, other Brooklyn Law School professors, as well as the school administration for contributing to the Law Review’s success. You have served as an inspiration to us, and it is thanks to your work and ideas that we have had the pleasure and privilege to play a part in the Law Review’s distinguished history.

Andrei Takhteyev
Editor-in-Chief, 2009-10
Introduction

Edward K. Cheng

It is my privilege to introduce this special issue of the Brooklyn Law Review in honor of my colleague and mentor, Margaret A. Berger. When Margaret announced her retirement from almost forty years of teaching at Brooklyn Law School, many of us were at a loss as to what to do. Naturally, as local tradition dictates, there would be a retirement dinner, complete with toasts from colleagues and family, speeches by former students, and the unveiling of a portrait. But for Margaret, the occasion seemed to demand something extra, something to acknowledge her remarkable contributions to the law of evidence and her eminent place in the field.

Out of that vaguely unsettled feeling, Larry Solan and I, in consultation with our dean, Joan Wexler, arrived at a “festschrift.” Using foreign phrases of course has its dangers, and throughout this process the Berger “Festschrift” has not infrequently been greeted with puzzled looks and tentative pronunciations. In concept though, everyone has perfectly understood not only what it is, but also why it is a perfect fit for Margaret’s retirement. A new collection of works on the law of evidence looking not only at the past, but also to the future—what better way to celebrate Margaret’s career?

Margaret’s career in the law of evidence has spanned the full breadth of the discipline. As a scholar, she has co-authored (with Judge Jack Weinstein) arguably the preeminent
treatise on the Federal Rules of Evidence, as well as written a number of influential law review articles. As a teacher, she has co-authored a leading evidence casebook, enlightened generations of students, and educated judges about the complexities of science with her now-famous Science for Judges conferences. As a lawyer, she has written amicus briefs to the Supreme Court and served on pathbreaking committees of the National Academy of Sciences. And as a member of the academic community, Margaret has embraced her role as part of the vanguard for women in the legal academy, and she has served as an important role model for young scholars female and male alike.

Befitting a career of such incredible scope, sixteen evidence scholars make their contributions in this festschrift issue. The word “festschrift” comes from the German, which can be literally translated as “party writing” or “festival writing” and this festschrift lives up to that name. As with most parties (good ones, at least), the sounds emanating from this festschrift are celebratory, boisterous, and more importantly, *polyphonous*. The articles run the gamut of modern evidence law.

The festschrift begins with three tributes. Judge Jack Weinstein is a natural tribute writer for any evidence scholar, but his tribute is a particularly fitting one in this case. After all, Margaret was not only Judge Weinstein’s first law clerk, but she has collaborated with him for years on both their treatise and casebook. Eleanor Swift writes a moving tribute to Margaret’s impact on women teaching in the field of evidence, and Larry Solan writes one as her long time colleague at Brooklyn Law School.

As might be expected given Margaret’s scholarly focus over the last two decades, scientific evidence articles comprise the lion’s share of this issue. Jennifer Mnookin, who delivered the Ira M. Belfer Lecture that preceded Margaret’s retirement dinner, considers the problems facing the forensic identification sciences. She devotes considerable effort addressing a key question left largely unexplored by the recent National Academy of Sciences report—how should courts respond to the problems found in forensic science?

The forensic themes of the Mnookin lecture surface in other contributions. Paul Giannelli, for example, chronicles Margaret’s involvement in the three biggest events in the recent history of forensic science—the Supreme Court’s decision in *Daubert v. Merrell Dow Pharmaceuticals*, the
acceptance of DNA evidence in the courtroom, and the recent National Academy of Sciences report. Ed Imwinkelried continues this vein by exploring Margaret’s role as amicus curiae in both Daubert and Kumho Tire v. Carmichael.

A different variation on the forensic theme appears in David Kaye’s contribution, which responds to a recent article by Jay Koehler and Michael Saks on the “individualization fallacy” in the forensic sciences. Kaye contends that they overstate their claim. Koehler and Saks in turn write a contribution in response.

In other scientific evidence pieces, an interesting thread emerges among several articles that highlights the tension between the general nature of science and the particularized nature of legal inquiry. As David Faigman notes in his contribution, this “evidentiary incommensurability” lies at the root of the tension between science and law, and he provides a useful taxonomy for navigating the terrain. Joe Sanders takes this dichotomy in a different direction, using it to explain the longstanding conundrum of why courts seemingly apply Daubert more strictly in civil cases than in criminal ones. I also rely heavily on the distinction in my contribution, which argues that scientific facts should be treated not like ordinary adjudicative facts but rather like foreign law.

Beyond scientific evidence, other evidentiary concerns and doctrines make appearances as well. Michael Risinger goes back to basics, arguing that reform of the evidentiary rules should strive to improve accuracy, as determined through a combination of empirical evidence and common sense. In contrast, Aviva Orenstein tackles a specific evidentiary bête noire, advocating that the rule governing past convictions for crimes of deceit under Rule 609(b) should be subject to a judicial balancing test like Rule 403.

Three contributions comment on the Confrontation Clause, motivated in part by Margaret’s pre-Crawford work on confrontation issues: her 1992 article discussing the Confrontation Clause implications of statements solicited by government actors, and her amicus briefs in Idaho v. Wright and Lilly v. Virginia. Bob Mosteller argues that informant testimony should be recorded in “draft form” to reveal the potential influence of government agents. Norman Abrams writes about the implications of Davis v. Washington’s “ongoing emergency” qualification to the Confrontation Clause for terrorism prosecutions and government attempts to gain information for intelligence purposes. Myrna Raeder discusses
the future of forfeiture doctrine in domestic violence cases in the wake of *Giles v. California*.

Finally, no festschrift in honor of Margaret could be complete without some words about teaching, and Roger Park kindly obliges with his reflections on “clickers.” With its often technical yet deterministic rules as well as its link to real-time courtroom objections, Evidence may be the most amenable among law school survey courses to the “clicker” revolution. Park offers one example of how to inject “clicker” technology into the classroom successfully.

* * *

So this festschrift is polyphonic indeed. And although detractors sometimes characterize the field of evidence as a narrow set of arcane rules—perhaps even a narrow set of *irrelevant* rules given the death of the trial—this celebration shows the field to be vibrant and wide-ranging. This result should come as no surprise. After all, the process of proof and the rules that govern it are not only fundamental to legal inquiry, but also necessarily reflect our deepest values.

A final word of thanks. Although the footnote at the beginning of this introduction acknowledges all of the people who made this festschrift possible, I purposely made one glaring omission, saving the best for last. I think it safe to say that all of the participants, as well as everyone else in the field of evidence, would gladly join me in thanking Margaret for dedicating her career to our field and helping make it what it is today. It is our great fortune to have her as a colleague and friend, and we look forward to her new contributions for many years to come.
TRIBUTES TO PROFESSOR
MARGARET A. BERGER

Tribute

The Honorable Jack B. Weinstein†

How fortunate are those of us whose lives have been enriched by Margaret Berger.

A superb classroom teacher of civil procedure, evidence, mass torts, and science and the law, she has provided the basis for professional success, a life of the intellect, and the ethical practice of law for thousands of her students.

For hundreds of state and federal judges, scientists, and government administrators of science-based programs, she has presented an extraordinary series of seminars and working sessions that have measurably improved the capacity of lawyers, scientists, and government officials to administer the law.

Colleagues in teaching and practicing lawyers and judges have been grateful for her influential casebooks, treatises, and articles.

To her many friends, she epitomizes the supportive, enchanting, engaged personal relationships that add sweetness to life even in its most trying hours.

And there is yet another dimension that is at the core of this uncommon woman. It is glimpsed from time to time by those who have had the pleasure of visiting her home. There she presides with warm elegance in her beautifully old-world furnished apartment in midtown Manhattan and her suburban house by the sea. Margaret, along with Mark, her late husband, a well known general practitioner, and their sons, Josh, a T.V. stage manager, and David, a boutique firm litigator, have often

† Senior District Judge, United States District Court for the Eastern District of New York.
invited students, colleagues, and friends to share their beautiful and bountiful table and cultured conversation.

It has been my privilege and honor to know Margaret Berger. She was my student (and thus my teacher) and my first law clerk. We have been coauthors and joint instructors. She has been a dear friend for over half a century. My bias in her favor is no disqualification from participating in this well deserved Berger Festschrift and dinner in her honor. As demonstrated by this outpouring of affection and by the many tributes to her and her Science for Judges programs in the sixteenth volume of the *Journal of Law and Policy*, legions will confirm that here partiality is mandated by truth.
Tribute

Lawrence M. Solan

Margaret Berger and I have been colleagues at Brooklyn Law School since I joined the faculty in 1996. In the beginning, I really didn’t know what to make of her. Some people are best described as “no-nonsense,” others as “warm and friendly,” still others as “brutally honest.” Margaret is all of those things among others, and sees no conflict among them. She is a complex and brilliant woman. I learned this about her fairly quickly, and we soon became close friends.

I know Margaret’s work well, and have even been fortunate enough to have written with her, which I hope to do again. Her work reveals a tightly argued, yet compassionate vision of the law. To take one example: Legal scholars have been concerned about the fact that there is not enough science in what passes as scientific evidence. Margaret has been part of this movement. But at the same time, Margaret has forged another, contrarian movement, whose theme is that the fetish of unattainable scientific certainty is being used by the courts to ensure that those injured by chemical and pharmaceutical products cannot recover for their injuries. That is because the courts require scientific proof of causation that is stricter than the scientific community can generally provide.

Margaret proposes changes in the substantive law, as well as the law of evidence and procedure, the latter two of which she regards as parts of the same package. She not only sees the big picture, but she is willing to paint a new one if that is what is needed.

For those who know how deeply Margaret has devoted herself to music, opera, and theater (for most of her life with her late husband, Mark), it should not be surprising that her work is as humanistic as it is intellectually crisp. I can’t say with scientific certainty that Margaret Berger is one of a kind,

---

† Associate Dean for Academic Affairs and Don Forchelli Professor of Law, Brooklyn Law School.
but I’ve been attempting to falsify that hypothesis for some time, and have come up empty so far.
Tribute

Eleanor Swift†

For women teaching in the field of evidence law, Professor Margaret Berger’s career shines as a beacon—a stellar teacher, a highly-respected scholar of scientific evidentiary issues, co-author of the most venerated of evidence casebooks, co-author of the foundational treatise on the Federal Rules of Evidence, consultant to courts, private and government commissions too numerous to mention, and Reporter to (among other august institutions) the Advisory Committee on the Federal Rules of Evidence. She has broken every glass ceiling in academia.

It could not have always been easy. Margaret entered the legal profession at a time when women were not hired by major New York law firms. She entered law teaching just as women were becoming a significant presence in law school student bodies, but not yet in law school faculties. And in the field of evidence law, male professors were iconic figures. Yet I wonder whether Margaret would ever admit how hard it was. She would probably give her characteristic “shrug” to underplay her own remarkable achievements.

These achievements paved the way for women who followed her into the academic profession and into the field of evidence law. It is easier not to be the first, even though Margaret always set such a high standard. Even her introductions are masterpieces, as those who heard her remarks about Judge Weinstein at the AALS Evidence Section Luncheon in 2008 well remember. And she paved the way with her own brand of fellowship as well as her own achievements. At every major Evidence conference, at which she was always an invited speaker, Margaret welcomed us. She shared her inquiring mind with us, immediately treating us as colleagues instead of newcomers.

It is her combination of fellowship and accomplishment that I have admired from near and afar for more than twenty

† Professor of Law, University of California, School of Law.

1065
years. I know this admiration is shared by countless other women (and men too) who have been similarly inspired by Professor Margaret Berger.
I. INTRODUCTION

Professor Margaret Berger is best known as a distinguished Evidence scholar on the subject of expert testimony. Her work includes, however, an important paper on the Confrontation Clause, published in 1992, titled, The Deconstitutionalization of the Confrontation Clause: A Proposal for a Prosecutorial Restraint Model, which was one of the earliest articles to foreshadow Justice Scalia’s majority opinion in the Supreme Court’s 2004 decision in Crawford v. Washington.

I dedicate this paper as a tribute to Margaret’s oeuvre on evidence and in celebration of that article, noting that she presciently anticipated the general change of direction by the Court, although Crawford took a different, related doctrinal
approach to the application and scope of the Confrontation Clause.

This paper will examine the implications of the 2004 Fourth Circuit compulsory process decision in United States v. Moussaoui, and also refer to its 2008 Second Circuit progeny, United States v. Paracha. It will use the doctrine of those cases, considered in light of the Supreme Court’s decisions in Crawford and Davis v. Washington, as the basis for an exercise to examine how certain types of confrontation and hearsay issues might be analyzed in future terrorism prosecutions in the federal courts—applications that neither Justice Scalia nor Professor Berger may have anticipated.

II. CRAWFORD V. WASHINGTON AND PROFESSOR BERGER’S PROPOSAL

It is, of course, familiar stuff to evidence scholars that Crawford dramatically shifted the Court’s approach to the Confrontation Clause, holding that “[t]estimonial statements of witnesses absent from trial” are admissible “only where the declarant is unavailable and only where the defendant has had a prior opportunity to cross-examine [the witness].” The key to this new doctrinal development is the notion of testimonial statements, which Justice Scalia, relying on and applying the history behind the Confrontation Clause, described in the following terms:

An accuser who makes a formal statement to government officers bears testimony in a sense that a person who makes a casual remark to an acquaintance does not. The constitutional text, like the history underlying the common-law right of confrontation, thus reflects an especially acute concern with a specific type of out-of-court statement.

Various formulations of this core class of “testimonial” statements exist, [e.g.]: “ex parte in-court testimony or its functional equivalent—that is, material such as affidavits, custodial examinations, prior testimony that the defendant was unable to cross-examine, or similar pretrial statements that declarants would reasonably expect to be used prosecutorially” . . . . These formulations all share a common

---

4 313 F. App’x. 347 (2d Cir. 2008). The circuit opinion, which was not certified for publication, treats the relevant issues summarily. For a fuller treatment, see the district court opinion, United States v. Paracha, No. 03-CR-1197(SHS), 2006 WL 12768 (S.D.N.Y. Jan. 3, 2006).
6 Crawford, 541 U.S. at 59.
nucleus and then define the Clause’s coverage at various levels of abstraction around it . . .

Statements taken by police officers in the course of interrogations are also testimonial under even a narrow standard. Police interrogations bear a striking resemblance to examinations by justices of the peace in England. The statements are not sworn testimony, but the absence of oath was not dispositive.7

Twelve years before Crawford was decided, Professor Berger advocated a related approach:

Hearsay statements procured by agents of the prosecution or police should therefore stand on a different footing than hearsay created without government intrusion. The Confrontation Clause should bar hearsay statements elicited by governmental agents unless the declarant is produced at trial or unless special procedures [which she later describes] . . . are followed.8

Both approaches would apply the Confrontation Clause to hearsay statements obtained by governmental agents for use at trial. The main difference appears to be that Justice Scalia uses a doctrinal category, “testimonial statements,” to characterize the type of hearsay covered by the Confrontation Clause while the emphasis under Professor Berger’s proposal is on the fact that the statement was procured by government agents and on government behavior. Rather than relying on a formulaic approach, she proposes one that would focus on the circumstances of the particular case and identify whether inappropriate forms of prosecutorial behavior were involved: Inter alia, she recommends that contemporaneous recordings be used in appropriate cases to ensure that the prosecutor or police did not pressure, induce, or manipulate the declarant into making the statement. She views the Confrontation Clause as a limitation on government action, as a way of keeping “the overwhelming prosecutorial powers of the government in check.”9

Use of the formulaic approach led Justice Scalia in Davis v. Washington to add a qualification to the definition of “testimonial,” and a key question for this paper is how broadly the Davis doctrine is to be interpreted. The very use of the definitional term, “testimonial,” thus carries with it

---

7 Id. at 51-52 (quoting Brief of Petitioner at 23, Crawford v. Washington, 541 U.S. 36 (2004) (No. 02-9410)).
8 Berger, supra note 1, at 561-62.
9 Id. at 562.
substantive implications. It is not clear whether Professor Berger would recognize a similar categorical limitation since governmental action would still be involved even in situations excluded from the coverage of the Confrontation Clause under the Scalia approach in *Davis*. If not, her approach would sweep more broadly than Justice Scalia’s and might bring more statements under the protection of the confrontation mantle.

III. THE IMPLICATIONS OF *DAVIS V. WASHINGTON*

As noted above, *Crawford* ruled that the introduction into evidence by the prosecution of out-of-court statements that are testimonial, where the declarant is unavailable and the statements had not been subject to prior cross-examination, violates the Confrontation Clause. *Davis v. Washington* qualified *Crawford* by treating as nontestimonial certain types of statements elicited by police questioning.\footnote{10} Recall that in describing the category of testimonial statements, Justice Scalia in *Crawford* stated: “Statements taken by police officers in the course of interrogations are . . . testimonial under even a narrow standard.”\footnote{11} We learn, however, from Justice Scalia’s opinion in *Davis v. Washington* about statements obtained through police questioning that are nontestimonial:

> Statements are nontestimonial when made in the course of police interrogation under circumstances objectively indicating that the primary purpose of the interrogation is to enable police assistance to meet an ongoing emergency. They are testimonial when the circumstances objectively indicate that there is no such ongoing emergency, and the primary purpose of the interrogation is to establish or prove past events potentially relevant to later criminal prosecution.

. . . .

. . . When we said in *Crawford* . . . that “interrogations by law enforcement officers fall squarely within [the] class” of testimonial hearsay, we had immediately in mind (for that was the case before us) interrogations solely directed at establishing the facts of a past crime, in order to identify (or provide evidence to convict) the perpetrator. The product of such interrogation, whether reduced to a writing signed by the declarant or embedded in the memory (and perhaps notes) of the interrogating officer, is testimonial. . . . A 911 call, on the other hand, and at least the initial interrogation conducted in connection with a 911 call, is ordinarily not designed

\footnote{10} *Davis*, 547 U.S. at 822.

\footnote{11} *Crawford*, 541 U.S. at 52.
primarily to “establish[or] prove[ ]” some past fact, but to describe current circumstances requiring police assistance.\textsuperscript{12}

In \textit{Davis}, Justice Scalia drew a distinction that had not surfaced in \textit{Crawford}—between testimonial and nontestimonial statements that result from police interrogation. The distinction rests principally on what was “the primary purpose of the interrogation.”\textsuperscript{13} If the interrogation is “solely directed at establishing the facts of a past crime, in order to . . . provide evidence to convict . . . the perpetrator,” it is testimonial.\textsuperscript{14} If its primary purpose was “to meet an ongoing emergency,” or was “designed primarily . . . to describe current circumstances requiring police assistance,” it is non-testimonial.\textsuperscript{15}

As we shall see further along in this paper, \textit{Crawford} and \textit{Davis} combined with \textit{Moussaoui} and \textit{Paracha} provide the basis for a line of argument that might be used to address confrontation concerns that would arise when the government tries to offer into evidence the results of the interrogation of suspected terrorists.

\section{THE DOCTRINE OF \textit{UNITED STATES V. MOUSSAOUI} AND \textit{UNITED STATES V. PARACHA}}

\subsection{Compulsory Process/Classified Information Issues}

Zacarias Moussaoui was indicted for a series of terrorism offenses including involvement in the conspiracy that led to the horrific acts on September 11, 2001. The 2004 decision of the Fourth Circuit U.S. Court of Appeals in \textit{United States v. Moussaoui},\textsuperscript{16} the final direct review opinion among numerous \textit{Moussaoui} decisions in the course of this prosecution, addressed compulsory process-classified information issues in the case.

\textit{United States v. Paracha}\textsuperscript{17} involved similar issues, though the charges in that case were based on identification document fraud and providing material support to al Qaeda in aid of terrorism. In \textit{Paracha}, the district court relied heavily on

\begin{thebibliography}{9}
\item Davis, 547 U.S. at 822, 826-27 (citations omitted).
\item Id. at 822.
\item Id. at 826.
\item Id. at 827-28.
\item 382 F.3d 453 (4th Cir. 2004), cert. denied, 544 U.S. 931 (2005). The opinion in volume 382 was a revised version of an earlier opinion, 365 F.3d 292 (4th Cir. 2004).
\end{thebibliography}
the Moussaoui opinion, and the circuit court summarily affirmed the lower court decision. The significance of these two Paracha opinions is not what they add substantively (though they do add slightly to the earlier Moussaoui case), but rather the fact that they are subsequent decisions that appear to indicate that, at least in a terrorism prosecution context, the Moussaoui doctrine that we examine in this paper is not “a derelict on the waters of the law.”

The Moussaoui Fourth Circuit decision, by happenstance, was handed down immediately after the Supreme Court issued its opinion in Crawford v. Washington. Accordingly, while Crawford is mentioned twice in Moussaoui, the references are brief. Neither the government’s nor the defendant’s brief in Moussaoui mentioned Crawford (since both briefs were prepared before the Crawford decision), although they do address some confrontation issues. Moussaoui is generally and correctly perceived as primarily involving, at a constitutional level, compulsory process issues; it is not a confrontation or hearsay decision. Yet, as discussed below, because of the way the court formulated its opinion, arguments regarding confrontation and hearsay issues in terrorism prosecutions can be derived from the case.

At the outset, two aspects of the Moussaoui opinion should be noted, which, while they make it more difficult to determine its meaning, do not prevent us from digging into the doctrine of the case. First, because the case involves classified information, the opinion is heavily redacted so that particular sentences are inconclusive and hard to decipher. (However, somewhat surprisingly, one can reasonably guess at some of the words that were deleted from the opinion.) Second, after the Supreme Court denied certiorari review of the circuit court’s opinion under consideration here, Moussaoui pleaded

---

18 United States v. Paracha, 313 F. App’x. 347, 351 (2d Cir. 2008).
19 Lambert v. California, 355 U.S. 225, 332 (1957) (Frankfurter, J., joined by Harlan, J. and Whittaker, J., dissenting). It should be kept in mind that Moussaoui is a Fourth Circuit opinion, and that Circuit has had the reputation of being one of the most conservative in the nation. See Neil A. Lewis, Obama’s Court Nominees Are Focus of Speculation, N.Y. TIMES, Mar. 11, 2009, at A19 (suggesting that because of the large number of potential appointments to that bench, President Obama has the possibility of turning the conservative Circuit quickly around). The fact that Moussaoui was followed by the Second Circuit in Paracha is therefore not without significance.
20 See Moussaoui, 382 F.3d at 461, 481.
21 See Brief of Appellee at 81-84, Moussaoui, 382 F.3d 453 (No. 03-4792); see also Brief for the United States at 48-60, Moussaoui, 382 F.3d 453 (No. 03-4792).
The circuit court opinion had contemplated further action in the district court: the judge was expected to apply the approach mandated by the circuit majority to the substitutions, but the guilty plea served, to an extent, to limit such action and also served to insulate the circuit court opinion from further direct review.

As it turned out, subsequently, Moussaoui attempted to withdraw his plea of guilty and in January 2010, a different circuit panel issued an opinion that revisited some of the relevant issues. This new opinion did not, however, change the essential conclusions of the first panel.

An important part of Moussaoui’s defense had been to show that he was not involved in the 9/11 conspiracy. One of the important issues in the case revolved around Moussaoui’s efforts to obtain testimony from three individuals, enemy combatant witnesses (the “ECWs”) at the time allegedly in U.S. custody abroad, who were alleged to be key al Qaeda members involved in the 9/11 conspiracy. Moussaoui asserted that these individuals could provide exculpatory testimony on his behalf—that he had not been involved in the 9/11 conspiracy. At the time, the government declined officially to acknowledge that these men were in custody, assuming that they were only for purposes of allowing the court to make a ruling whether to issue a subpoena ad testificandum. The government refused, on grounds of national security, to produce the ECWs in person or to make them available so their depositions could be taken.

In an effort to resolve the impasse thereby created, a compromise was attempted by the Fourth Circuit Court of Appeals—to prepare substitutions for the testimony of these three witnesses. While the case did not directly bring into play the Classified Information Procedures Act (“CIPA”), because live witness testimony was at issue, both the district and circuit courts relied heavily on the CIPA approach for dealing with classified evidence that the government does not wish to

---

22 United States v. Moussaoui, 591 F.3d 263, 271 (4th Cir. 2010).
23 See id. at 271-72, 276 (“The substituted statements of ... several other terrorists were ... admitted as evidence during the sentencing proceedings.”).
24 See id. at 284-85. In considering Moussaoui’s attack on his conviction and attempted withdrawal of his plea of guilty, the January 2010 opinion reiterated the previous panel’s justifications for treating the substitutions as reliable. See id.
25 Id. at 271, 287.
26 Id. at 271.
produce because of concerns about the need for secrecy and confidentiality. An important tool under CIPA for dealing with such evidence is to fashion substitutions, usually summaries, which do not disclose any of the information that the government wishes to keep confidential, but provide a defendant with enough evidence to enable him/her to make his/her defenses.

The district court had concluded that the substitutions that had been offered by the government did not and could not adequately protect the defendant’s right to defend himself. Reversing the lower court on this point, the circuit panel concluded that, with more work and participation by the district judge and the parties in the preparation of the substitutions, both the government’s national security interests and the compulsory process and right-to-defend interests of the defendant could be adequately protected.

The substitutions approved in principal by the court had unique features. They were summaries, but of what? The witnesses had not given depositions. Rather, detailed statements (1) from the putative three ECWs, labeled by the court as A, B and C, had been obtained, apparently through interrogation, by government agents seeking intelligence that could be used in preventing future terrorist actions and apprehending other terrorists. (We do not address in this paper any issues arising out of claims that the statements at issue might have been obtained by government agents or their surrogates through interrogation involving torture or other forms of coercion.)

The statements thus obtained from A, B, and C had then been “recorded” in “highly classified” reports (2) which had been prepared for “use in the military and intelligence communities; they were not prepared with this litigation” in mind. Portions of the reports were then “excerpted and set forth in documents prepared for purposes of this litigation”

---

29 See, e.g., Moussaoui, 382 F.3d at 471-72, 476-77.
30 See Moussaoui, 591 F.3d at 284.
31 Id.
32 Moussaoui, 382 F.3d at 458.
33 See id. at 458 n.5.
34 Id.
which were labeled summaries (3). The summaries were provided to defense counsel who had security clearance.

Finally, the substitutions (4) to be offered into evidence were prepared from the summaries. These largely consisted of the summaries of statements that had been obtained over the course of several months but were not identical to these summaries since in preparing the substitutions, the government had reorganized the information.

It is apparent from the foregoing that there were multiple levels of hearsay involved in the substitutions that were being considered by the court, with A, B, and C being the declarants at level 1, supra, and the declarants involved in the preparation of the documents labeled as levels 2-4 being unnamed government agents or employees.

The district court had “deemed the substitutions inherently inadequate because the . . . reports, from which the substitutions were ultimately derived, were unreliable.” In part this was because “it cannot be determined whether the . . . reports accurately reflect the witnesses’ statements . . . . The [district] court further commented that the lack of quotation marks in the . . . reports made it impossible to determine whether a given statement is a verbatim recording . . . .”

Responding to the district court’s concerns, the circuit court stated:

The answer to the concerns of the district court regarding the accuracy of the [Redacted] reports is that those who are [Redacted] [ed. interrogating (?)] the witnesses have a profound interest in obtaining accurate information from the witnesses and in reporting that information accurately to those who can use it to prevent acts of terrorism and to capture other al Qaeda operatives. These considerations provide sufficient indicia of reliability to alleviate the concerns of the district court.

In the foregoing passage, the circuit court addressed two issues: first, whether the information obtained from A, B, and C was reliable, which it dealt with by stating that the questioners had “a profound interest in obtaining accurate

---

35 The district court noted “that it had been impressed with the accuracy of the summaries.” See id. at 478 n.30 (internal quotation marks omitted).
36 See id. at 478-79.
37 Id. at 478.
38 Id.
39 Id.
information from the witnesses’; and second, the accuracy of
the recording of those statements in the reports, to which it
responded with the statement, “[they] have a profound
interest . . . in reporting that information accurately.”

Relying on the fact that the interrogators’ motivation
was to obtain accurate statements from A, B, and C for
intelligence purposes (and, implicitly, that they were experts in
this activity) and that the documents recording those
statements were prepared in aid of a similar purpose, the court
concluded that with adequate work by the district court and
the parties, the substitutions could be prepared in a form that
was reliable, and in that form the defendant could introduce
them into evidence. Finally, the court ruled that the defendant
was entitled to have the fact that the substitutions are reliable
communicated to the jury via an appropriate instruction.

The court did not expressly parse the multiple levels of
hearsay involved in this material, nor did it expressly identify
the hearsay element that presented the greatest challenge to
the court’s overall assessment that the substitutions were
reliable, namely the reliability of the original statements by the
declarants. The court did briefly further address reliability of
this first level, however, when it made the following statement:

To the contrary, we are even more persuaded that the [Redacted]
process is carefully designed to elicit truthful and accurate
information from the witnesses.

Here the court seemed to introduce an additional idea to the
argument based upon the government agents’ motivation to
obtain accurate information: the very process [of interrogation
(?)] was carefully designed to obtain truthful and accurate
information.

In discussing the reliability of these out-of-court
statements, the circuit court did not expressly mention the
hearsay rules or discuss the issues in traditional hearsay
terms. This is only somewhat surprising. Although the
substitutions were hearsay and the court was addressing a
traditional hearsay concern—namely, reliability—the court
here was dealing with a compulsory process issue, not a
confrontation/hearsay issue. The evidence sought in this case

40 Id.
41 Id.
42 Id. at 478-79.
43 Id. at 478 n.31.
was hearsay, in the context, to be introduced by the defendant, not by the prosecution, and ordinarily a party cannot (and usually does not) complain about hearsay weaknesses of evidence that he/she has offered or seeks to introduce. Further, the opponent of the evidence to be offered, i.e., the prosecutor, was not likely to raise a hearsay objection because the court’s approach held promise of resolving the difficult compulsory process/classified information issues in the case that might otherwise have turned out to be an obstacle to obtaining a conviction.

Although it was the defendant who wished to introduce the evidence in question, he was not totally in control of the form of the evidence that might be presented. So, this was not a typical situation in which a party may not be heard to complain about hearsay that it offered into evidence.

United States v. Paracha involved a similar context and set of issues and the court relied on the Moussaoui approach in addressing the matter. Without calling any special attention to it, the district court in Paracha included in the jury instructions an additional ground that bore on the circumstantial reliability of the first level of hearsay in that case:

The failure [on the part of the witnesses/interrogtees] to provide truthful information would be detrimental to any relationship to the United States government by the witnesses.\(^44\)

Like the Moussaoui court, the district court in Paracha did not, however, address the issues raised by these documents in terms of traditional hearsay categories.

By giving instructions in Moussaoui and Paracha that the exculpatory\(^45\) statements contained in the substitutions were reliable, both courts figuratively leaned over backwards in favor of the defendant. Otherwise, absent a resolution of the impasse, the government might have been faced with the

\(^{44}\) United States v. Paracha, No. 03-CR-1197(SHS), 2006 WL 12768, at *15 (S.D.N.Y. Jan. 3, 2006). The sentence was probably included in the instructions because Paracha’s counsel had requested an instruction, which the court rejected, to the effect that the witnesses/declarants in question were providing assistance to the U.S. government. See id. at *14 n.2. Instead the court gave the indicated instruction which on its face seemed to suggest that the witnesses had some special relationship to the government, which would motivate them to tell the truth.

\(^{45}\) See Moussaoui, 382 F.3d at 479-82; Paracha, 2006 WL 12768, at *14-16. Note that in Moussaoui, the court conceded that not all of the statements in the substitutions were exculpatory. See Moussaoui, 382 F.3d at 473. The government had argued that a number of the statements in fact incriminated the defendant. See id.
prospect of a court-imposed sanction dismissing the prosecution. In assessing the courts’ actions here, it seems fair to conclude that in each of the cases, the court was trying to be responsive to the concerns about whether the defendant’s right to make his defenses would be infringed if neither the witnesses requested, nor their depositions, were produced.

What is worth emphasizing, however, is not why both courts did what they did, but rather, what they did. To ensure that each defendant’s right to defend himself was not infringed, the courts gave each defendant the benefit of an instruction that the substitutions and the exculpatory statements contained therein were obtained under circumstances that indicated that the statements were reliable.

It is the fact that both courts concluded that a) the statements were obtained for terrorism intelligence purposes (“in aid of the pursuit of terrorists and prevention of acts of terrorism”), and b) that in the circumstances the statements were reliable, which makes it possible to think about deriving arguments regarding confrontation and hearsay issues from these decisions. We discuss these lines of argument in the next section.

B. Addressing Possible Confrontation and Hearsay Issues in a Hypothetical Federal Prosecution of a Gitmo Detainee

1. A Hypothetical Situation

Suppose that the issue of the admissibility of Moussaoui-type substitutions were to arise in a future federal court terrorism prosecution. Assume, however, that the

---


In the military commission proceedings, the statutory rules for dealing with hearsay and coerced confessions are different from the rules applied in civilian courts. (In light of the decision in Boumediene v. Bush, 128 S. Ct. 2229, 2262 (2008) (holding that enemy combatants tried in military commissions could not be deprived access to the federal writ of habeas corpus), a lurking set of questions is whether and to what extent specific constitutional protections will be available to detainees being prosecuted in military commission trials. Specifically, will they be protected by the Confrontation Clause? If so, the discussion in the text regarding confrontation clause
substitutions contain statements that incriminate the defendant and resulted from the interrogation of detainees whom the government is unable to produce. Assume further that this time it is the prosecutor who offers the substitutions containing these statements into evidence against the defendant. Are there arguments derived from the opinions in Moussaoui and Paracha doctrine that can be used in addressing this issue?

2. Implications of Crawford and Davis for the Confrontation Issue

To understand the potential significance of the Moussaoui and Paracha doctrine applied in confrontation and hearsay contexts, we briefly return to the Supreme Court's decisions in Crawford v. Washington and Davis v. Washington. Recall that Davis qualified Crawford by ruling that not all statements obtained from police questioning are testimonial. We quote again from the text, supra, that describes the}

---

issues would be relevant as well to military commission proceedings. Cf. Al-Bihani v. Obama, 590 F.3d 866, 879 (D.C. Cir. 2010) (holding that Confrontation Clause does not apply in habeas corpus suits, applies only in criminal cases). The discussion in the text regarding a hypothetical prosecution assumes that it occurs in a federal district court. Given the impending prosecutions, the issues addressed in the hypothetical situation could turn out to be not quite so hypothetical.

A question may be raised as to whether this hypothetical scenario is realistic. Is the government likely to attempt to introduce against one of the Guantanamo detainees on trial for terrorist offenses a statement implicating him obtained through government interrogation from one of his alleged accomplices? Attempting to introduce into evidence against a defendant an accomplice's statement obtained by the police is not an uncommon practice.

Of course, if the witness/declarant is in government custody, under the confrontation doctrine applicable in the federal courts, the government would have an obligation to produce him if he was available, rather than use his hearsay statement. One can imagine situations, however, in which the witness/declarant is unavailable. He might, for example, as some terrorist detainees have done, refuse to cooperate with the proceedings, or he may be unavailable for some other reason.

Whether in a real case, the use of the Moussaoui-type substitutions would be involved if the government were attempting to offer a detainee's statement into evidence might depend on the circumstances. If the detainee-declarant is unavailable, the government might offer into evidence the fruits of its interrogation of that individual, which might be in the form of intelligence reports gleaned from interrogations conducted for intelligence purposes. If these intelligence reports contained sensitive information, there might be a need to summarize them and prepare the same kind of substitutions used in the Moussaoui case. On the other hand, if there were no need to classify information in the report, preparing summaries and substitutions would not be necessary. Note that at the time, in Moussaoui, the government was not even willing to admit that it had the witness/declarants in custody.
grounds for distinguishing between testimonial and nontestimonial statements:

The distinction rests principally on what was “the primary purpose of the interrogation.” If the interrogation is “solely directed at establishing the facts of a past crime, in order to . . . provide evidence to convict . . . the perpetrator,” it is testimonial. If its primary purpose was “to meet an ongoing emergency,” or was “designed primarily . . . to describe current circumstances requiring police assistance,” it is non-testimonial.48

Given the distinction drawn by Davis, an argument can be made, based on Moussaoui-Paracha, for addressing the confrontation concerns posed in our hypothetical situation. Those two courts indicated that in those cases the purpose of the interrogation of the suspected terrorists (the ECWs) by government agents was to provide intelligence to prevent other catastrophic terrorism events from occurring and to apprehend other dangerous terrorists. The two cases relied on the fact that the government interrogators had a purpose other than trying to obtain evidence to convict the suspected terrorists. Generally, the existence of such an alternative purpose avoids many of the concerns that lie behind the Crawford doctrine. But, as mentioned below, Davis is subject to varying interpretations. Does the alternative purpose of the interrogation fall within the specific terms of the Davis doctrine?

While this alternative purpose may not have involved an “ongoing emergency,” it is arguable that it was intended to obtain information regarding “current circumstances requiring police assistance [action?].” Preventing terrorism events that may soon happen may not amount to an “ongoing emergency” of the very immediate type involved in Davis (although even that point is debatable), but it has some similar characteristics. Certainly, the government can argue that the purpose of the interrogation was geared to an urgent need to obtain information quickly in order to prevent the occurrence of serious terrorism events that could occur at almost any time. If the interrogations occurred in the immediate aftermath of 9/11, the argument would be bolstered. Most importantly, it can be emphasized that the primary purpose of the interrogation was not to establish the facts of a past crime.

48 See supra text accompanying notes 13-15.
What is the likelihood that prosecutors in future cases would be successful in using this application of Moussaoui-Paracha as a basis for responding to the serious confrontation concerns posed in our hypothetical situation? It has been suggested by scholars that Justice Scalia’s opinion in Davis leaves room for differing interpretations of the kinds of police questioning that fall under the nontestimonial label; that the phrases used, such as “primary purpose” and “ongoing emergency” are “extremely ambiguous” and the ambiguity enables the judges to manipulate the concepts. Some scholars have focused on the requirement of an ongoing emergency. Still others suggest that the Crawford-Davis testimonial test is “almost arbitrary in its result” and no better than the earlier reliability approach.

If the courts were prepared to invoke the Moussaoui doctrine in this way, they would want to be certain that the result would not open the door to application of the doctrine in ordinary criminal prosecution contexts. Suppose, for example, that in an ordinary prosecution the police claim the purpose of the interrogation is to discover information that would be helpful in apprehending one of the defendant’s partners, not to help convict the defendant. One can imagine that the courts would formulate some kind of limiting principle(s) to avoid such a result—e.g., that the alternative purpose must not be ad hoc but must amount to a sustained and continuing intelligence-gathering effort, or as in Davis, must be the “primary purpose.” In Davis, the emergency nature of the situation may also be viewed as the Court’s way of limiting the

50 See, e.g., Roger Kirst, Confrontation Rules after Davis v. Washington, 15 J.L. & POL’Y 635, 641-44 (2007). Professor Kirst also details in his article how in the wake of the Davis decision the Supreme Court disposed of cases applying the Davis criteria.
51 Raeder, supra note 49, at 776. While ambiguous general terms were used in the Davis opinion, Justice Scalia did, however, also mention specific facts that can be argued to limit the scope of application of the doctrine. For example, the fact that the interrogations took place in a calm setting and the statements were made in response to a series of questions with the “officer-interrogator taping and making notes of . . . [the] answers” were mentioned as factors in favor of treating the statements as testimonial. Davis, 547 U.S. at 827. Invocation of such specific facts from the Davis situation, would provide a basis for arguing that the Davis exception should not be applicable in our hypothetical situation. See Raeder, supra note 49, at 775-76.
scope of the exception. In our hypothetical, the exigency and urgency of finding other terrorists before they perpetrate acts of catastrophic terrorism may be viewed as a similar type of limitation.\(^5\)

One might reasonably expect that the rationale of Moussaoui-Paracha would be invoked by the government to address the confrontation issue thus raised. While, given the arguments that can be made on both sides of the testimonial-nontestimonial issue, invocation of Moussaoui-Paracha would not guarantee that the government would prevail, it does provide a line of plausible, nonfrivolous arguments on the confrontation issue. In a prosecution of a serious terrorist defendant, such an argument would have a reasonable chance of success.

3. Addressing the Hearsay Issues in the Hypothetical Prosecution

If the government were successful on the confrontation issue, there would be a further set of questions to consider. As Justice Scalia wrote in Crawford, "Where nontestimonial hearsay is at issue, it is wholly consistent with the Framers’ design to afford the States flexibility in their development of hearsay law . . . ."\(^5\) If a court were to conclude that the statements at issue in our hypothetical situation are nontestimonial under Davis and therefore not inadmissible under the Confrontation Clause, it would be necessary to address the question whether the statements are nevertheless inadmissible under the jurisdiction’s rules governing hearsay.

The hypothetical case assumes a prosecution in a federal court; accordingly, the Federal Rules of Evidence would be applicable. Here, too, Moussaoui-Paracha appears to provide a basis for developing a line of arguments to address the

\(^5\) Compare the following: On March 1, 2010, the Supreme Court granted certiorari on the question:

whether preliminary inquiries of a wounded citizen concerning the perpetrator and the circumstances of his shooting are nontestimonial because they were “made under circumstances objectively indicating that the primary purpose of the interrogation is to enable police assistance to meet an ongoing emergency,” that emergency including not only aid to a wounded victim, but also the prompt identification and apprehension of an apparently violent and dangerous individual?


hearsay issues. We have previously described the several levels of hearsay involved in the Moussaoui case, and the rationales articulated by the Fourth Circuit in support of the reliability of all of those hearsay levels. We also identified the most serious hearsay weakness in the multiple levels of out-of-court statements—the fact that the original statements resulting from the interrogation do not appear to have any special indicia of reliability. How would the hearsay issues thus raised play out under the Federal Rules of Evidence?

a. Qualifying Levels 2-4 Under the Federal Rules

In addressing the admissibility of levels 2-4 in our hypothetical situation, conceivably, the prosecutor might invoke Federal Rule of Evidence 807, the Residual Exception, which deals with “[a] statement not specifically covered by Rule 803 or 804, but having equivalent circumstantial guarantees of trustworthiness” and then rely on the type of arguments made in Moussaoui regarding the reliability of the documents at issue there.\(^{55}\)

More likely, however, the prosecution would first argue that all of the levels of hearsay except the first level, i.e., the interrogatee’s statements, fall under the hearsay exception set forth in Rule 803(8), Public Records and Reports, which provides:

> Records, reports, statements, or data compilations, in any form, of public offices or agencies, setting forth . . . (B) matters observed pursuant to duty imposed by law as to which matters there was a duty to report, excluding, however, in criminal cases matters observed by police officers and other law enforcement personnel . . . unless the sources of information or other circumstances indicate lack of trustworthiness.\(^{57}\)

Recall that the materials in question in our hypothetical—i.e., the substitutions—are based upon 1) the statements made by the interrogatee, 2) the recording of those statements, in direct or indirect form, in reports, 3) the summaries of those statements and 4) the reorganizing of that material to make up the final form of the substitutions. An

\(^{55}\) Fed. R. Evid. 807.

\(^{56}\) There are additional requirements in Rule 807, for example, that the evidence is “more probative on the point than any evidence which the proponent can reasonably procure,” but we focus here on the equivalent trustworthiness element. Id.

\(^{57}\) Fed. R. Evid. 803(8).
argument can plausibly be made that levels 2) to 4) of the stages in the preparation of these substitutions amount to the preparation of reports by public officials under a duty to report within the meaning of Rule 803(8). The key question would then be whether they fall within the exclusion in that section applicable to criminal cases, for matters observed by police officers and other law enforcement personnel.

An often-cited case on the exclusion for police officer observations is United States v. Quezada which addressed the exclusion clause in the following terms:

While some courts have inflexibly applied the Rule 803(8)(B) proscription to all law enforcement records in criminal cases, . . . we are not persuaded that such a narrow application of the rule is warranted here. The law enforcement exception in Rule 803(8)(B) is based in part on the presumed unreliability of observations made by law enforcement officials at the scene of a crime, or in the course of investigating a crime: ostensibly, the reason for this exclusion is that observation by police officers at the scene of the crime or the apprehension of the defendant were not as reliable as observations by public officials in other cases because of the adversarial nature of the confrontation between the police and the defendant in criminal cases.

. . . .

. . . This circuit has recognized that Rule 803(8) is designed to permit the admission into evidence of public records prepared for purposes independent of specific litigation. . . . In the case of documents recording routine, objective observations, made as part of the everyday function of the preparing official or agency, the factors likely to cloud the perception of an official engaged in the more traditional law enforcement functions of observation and investigation of crime are simply not present. Due to the lack of any motivation on the part of the recording official to do other than mechanically register an unambiguous factual matter . . . , such records are, like other public documents, inherently reliable.

Quezada thus reasoned that the basis for the exclusion was that police officer observations at the scene are not reliable whereas other types of public official observations that have inherent reliability are not within the exclusion, such as records prepared for purposes independent of specific litigation—situations where there is a lack of any motivation

58 754 F.2d 1190 (5th Cir. 1985).
59 Id. at 1193-94 (emphasis added) (citations omitted) (internal quotation marks omitted).
on the part of the official to do other than mechanically register unambiguous factual matters.

It is arguable that the Moussaoui alternative purpose—the motivation to obtain accurate intelligence that could be acted upon by agents in the field in order to prevent future terrorist actions, not to gather evidence to be used in litigation—provides comparable grounds for concluding that the recordings and reports of the statements contained in the substitutions are “inherently reliable.”

Thus, some of the same arguments for treating the substitutions as nontestimonial under the Crawford confrontation doctrine, can be invoked to support qualifying all but the first level of hearsay in the substitutions under Rule 803(8) of the Federal Rules of Evidence, and for not applying the criminal case/matters-observed-by-police exclusion in that rule.

However, the ultimate admissibility of levels 2-4 is dependent under Rule 803(8) on whether “the sources of information . . . indicate a lack of trustworthiness.” If the statements made by the detainee-declarants as summarized in the substitutions are deemed unreliable and “indicate a lack of trustworthiness,” nothing of the substitutions—that is, levels 2-4—is admissible.

60 Fed. R. Evid. 803(8).

61 The Moussaoui arguments discussed in the text up to this point can be offered in support of the reliability of the reporting, summarizing and recordings actions of the government agents and to respond to any questions raised with respect to levels 2-4 under the last clause of Rule 803(8), “unless the sources of information or other circumstances indicate a lack of trustworthiness,” Fed. R. Evid. 803(8), but the question of the reliability of the first level statements remains and must be separately treated.

In Crawford, Justice Scalia stated that business records, which are analogous to public records are “by their nature . . . not testimonial.” Crawford v. Washington, 541 U.S. 36, 56 (2004); see also, e.g., United States v. Rueda-Rivera, 396 F.3d 678, 680 (5th Cir. 2005) (per curiam) (holding items in an immigration file akin to business records are non-testimonial in nature).

Regarding the question of whether the business records exception in the Fed. R. Evid. 803(6), Records of Regularly Conducted Activity, which does not have a criminal case/matters-observed-by-police exclusion might be used in lieu of Fed. R. Evid. 803(8), see United States v. Oates, 560 F.2d 45, 78 (2d Cir. 1977). Numerous cases have interpreted and limited Oates. These authorities are collected in Jack B. Weinstein, John H. Mansfield, Norman Abrams, Margaret A. Berger, Evidence—Cases and Materials 710-14 (9th ed. 1997).
b. Can the Interrogates' Statements (Level 1) Be Qualified Under the Federal Rules?

i. Intrinsic Circumstantial Reliability

As previously mentioned, the first level of hearsay, the interrogatee's statements to the government agents in response to questioning, presents the most problematic of the hearsay issues raised in the hypothetical situation.

The hearsay issue thus posed falls within the general category of a declarant's statements resulting from police interrogation being offered into evidence by the prosecution in the trial of his alleged accomplice. Efforts have often been made in the past, relying on the declaration against penal interest exception, to introduce such statements into evidence. It has been argued that such statements have intrinsic reliability that comes from being against the penal interest of the declarant. The asserted “against interest” features of a statement that inculpates a co-conspirator are, however, a complicated subject that has been addressed by scholars; and the Supreme Court, in a series of cases culminating in Lilly v. Virginia, usually ruled against the admissibility of such statements.

Conceivably, there might be some other type of special circumstances that would be suggestive of intrinsic circumstantial reliability of interrogation statements. A possible example of such a special circumstance might be the kind of facts underlying the aforementioned instruction given in the Paracha case, which suggested that the declarants had a motive to tell the truth arising out of the fact that they had developed a relationship with the government interrogators which they were interested in maintaining. The implication was that to lie to the interrogators would put that relationship at risk.

---


64 See supra note 44, and accompanying text.
This type of rationale is a weak ground for a claim of circumstantial reliability. Would it mean, for example, that the hearsay statements of an informer who works regularly with the police should be viewed as having intrinsic reliability because of his desire to maintain the relationship and whatever benefits accrue to him therefrom?

Whatever the merits of this specific type of claim, the Paracha instruction is an example of a claim of reliability based upon specific facts for a hearsay statement that does not fall under a specific exception. To be independently qualified, the statement would have to meet the standards of Rule 807, the Residual Exception, that is, circumstantial guarantees of trustworthiness equivalent to those under Rule 803 or 804.\footnote{65}

We also learn from Lilly v. Virginia, however, (which was decided under pre-Crawford confrontation-trustworthiness standards), that under the then-applicable constitutional doctrine, the requirements for meeting residual trustworthiness standards in a case involving accomplice hearsay resulting from government interrogation and incriminating a criminal defendant are likely to be very difficult to meet:

It is clear that our cases consistently have viewed an accomplice’s statements that shift or spread the blame to a criminal defendant as falling outside the realm of those “hearsay exception[s] [that are] so trustworthy that adversarial testing can be expected to add little to [the statements’] reliability.” . . . The decisive fact, which we make explicit today, is that accomplices’ confessions that incriminate a criminal defendant are not within a firmly rooted exception to the hearsay rule as that concept has been defined in our Confrontation Clause jurisprudence.

. . . .

The residual “trustworthiness” test credits the axiom that a rigid application of the Clause’s standard for admissibility might in an exceptional case exclude a statement of an unavailable witness that is incontestably probative, competent, and reliable, yet nonetheless outside of any firmly rooted hearsay exception . . . . When a court can be confident—as in the context of hearsay falling within a firmly rooted exception—that “the declarant’s truthfulness is so clear from the surrounding circumstances that the test of cross-examination would be of marginal utility,” the Sixth Amendment’s

\footnote{65} The claim of reliability would also be relevant to the issue regarding levels 2-4 under Fed. R. Evid. 803(8) applying the clause, “unless the sources of information . . . indicate lack of trustworthiness.” See supra notes 60-61 and accompanying text.
residual “trustworthiness” test allows the admission of the declarant’s statements.

It is highly unlikely that the presumptive unreliability that attaches to accomplices’ confessions that shift or spread blame can be effectively rebutted when the statements are given under conditions that implicate the core concerns of the old ex parte affidavit practice—that is, when the government is involved in the statements’ production, and when the statements describe past events and have not been subjected to adversarial testing.\footnote{Lilly, 527 U.S. at 133-34, 136-37 (alterations in original) (citations omitted).}

Although the Court’s expressed view of the “presumptive unreliability that attaches to accomplices’ confessions that shift or spread blame,” was written in a setting involving a constitutional issue under pre-Crawford standards, it would seem to bear generally on questions regarding the unreliability of such statements. It suggests that invocation of the Moussaoui-Paracha rationale based on a claim of intrinsic reliability because of special circumstances relating to the declarant’s statements would be faced with a strong presumption of unreliability.

ii. Extrinsic Circumstantial Reliability

A closer examination is also warranted of the specific elements underlying the Moussaoui court’s conclusion that the first level statements are sufficiently reliable based on extrinsic factors, that is, not arising from the intrinsic nature and content of the statement or motive of the declarant, but rather from the motivation and methods of the government agents who obtained the statements from the ECWs. The question is whether these factors, which were relied upon by the court in Moussaoui, are sufficient to establish circumstantial trustworthiness for purposes of Rule 807.\footnote{That is, whether the motivation of the interrogators to obtain accurate information (plus the implied appeal to the expertise of the interrogators) and the fact that the process of interrogation is “well designed” for such a purpose, are in combination sufficient circumstantial guarantees of the reliability of the statements. See \textit{Fed. R. Evid.} 807.}

The motivation and methods argument advanced by the court in Moussaoui can be broken down into three parts. The first part takes the following form: Ordinary police interrogators are motivated to obtain convictions—whether of
the person being interrogated or other persons. That motivation may shape how they question and the direction of their questioning. Government intelligence agents, such as those who interrogated the ECWs, are motivated to find other terrorists and prevent terrorist actions. They need completely accurate information for this purpose. They have no motivation to shape the direction of the information they obtain other than in ways designed to ensure that totally accurate information is obtained that will help them find terrorists and prevent acts of terror. However, it is a long leap to conclude that, because the interrogators are motivated to obtain accurate information they do indeed obtain accurate information from the persons whom they have interrogated.\footnote{There is a different tack that might be taken in addressing the Moussaoui court's conclusion that the statements of the ECWs were reliable—namely, introducing nonhearsay evidence regarding the reliability of the statements in an effort to meet the reliability threshold of Fed. R. Evid. 807. The court's comments suggested that the government agents viewed these statements as reliable; the implication was also that the government used and acted upon the information from the ECWs in the work done by government agents in the field to apprehend terrorists and prevent acts of terrorism. Suppose the question is whether testimony regarding such government actions should be admissible to prove the reliability of the statements made by the ECWs—that is, offering the government's actions to prove the belief of the government to prove the existence of that fact believed, i.e. the reliability of the statements. That tack is similar to what was done in the classic English evidence-hearsay case, Wright v. Tatham, (1837) 112 Eng. Rep. 488, 494-95 (Exch. Ch.). In that case, of course, a number of the justices concluded that the evidence in question was hearsay. Id. at 516-17. The definition of hearsay used in the Federal Rules of Evidence, however, excludes “from the operation of the hearsay rule all evidence of conduct, verbal or nonverbal, not intended as an assertion.” Fed. R. Evid. 801(a) advisory committee's note. Accordingly, the actions of the government in this regard would not be treated as hearsay under the Federal Rules. Cf. Cal. Evid. Code § 1200 cmt. (West 2009) (“[T]here is frequently a guarantee of the trustworthiness of the inference to be drawn from . . . nonassertive conduct because the actor has based his actions on the correctness of his belief, i.e., his actions speak louder than words.”).}

The second part of the argument which can be inferred from the Moussaoui opinion assumes that the interrogators have special expertise. What kind of expertise? We assume the claim is that the interrogators are persons of high intelligence, sensitivity and psychological insight who are able to make judgments about people, the logic of their stories, and to discern factual inconsistencies and flaws in the stories being told; that they also have the ability to make judgments about personalities, mannerisms, candor, dissembling, disingenuousness and the like, that make them specially capable of distinguishing truth-telling from false stories or
stories that have some falsity in them. Most judges are likely to be skeptical about such claims of expertise, especially about whether they are strong enough to support a finding of reliability.

The third part of the argument is that the process of questioning is “well designed” to obtain accurate information. What was the court referring to here? At the time, the government was not willing to acknowledge that it had the suspects in custody or that it had been interrogating them. Accordingly, there was no need to specify the methods used in interrogating them. We assume, however, that if special methods were being used in such interrogations, the government would try to keep the methods confidential and protected by the mantle of classified information. So we are faced with a factual claim that we are not able to assess because the facts relating to it are not available or likely to become available. Of course, such facts might be made available to the court in camera, and the court would be able to make a judgment (as the Moussaoui court apparently did) as to whether the nature of the process of interrogation assured that the statements obtained were sufficiently reliable.

In the absence of more specific information, we can only make some very general and highly speculative observations about what might be the design of a process of interrogation that would assure sufficient reliability. If the process involved the use of any form of coercion, apart from the due process and related claims that might be raised, the very use of coercion would tend to cast doubt on the reliability of the statements made. Or suppose the process involved the use of lie detectors, and that is the basis for the court’s judgment of sufficient reliability? Again, a Pandora’s Box of issues would be opened by such a claim. Suppose that drugs were used in interrogating the suspects? Again, due process issues would arise out of such a process. Or suppose that the government used some type of special psychological or other techniques, not involving any prohibited coercion, in eliciting the statements from the suspects?

The assumption thus seems to be that the persons who conduct the CIA (and other similar agencies) interrogations of terrorist suspects are much better at their job than police who conduct ordinary crime interrogations. We do not have enough information to make a judgment about the validity of such a claim.

See Steven Kleinman and Matthew Alexander, Op-Ed, Try a Little Tenderness, N.Y. TIMES, March 11, 2009, at A31. The authors of this article describe themselves as “military interrogators” who have questioned “hundreds of prisoners.”
It is impossible to evaluate the soundness of a claim of sufficient circumstantial reliability without knowing the specific facts underlying the claim. Accordingly, the ultimate resolution of the claim of sufficient reliability based on the methods of interrogation must here be left hanging in the air. It would be very troubling, however, if in connection with such an issue classified information is provided to the judge regarding the interrogation techniques used—information not available to the public—and the judge bases a ruling in favor of admissibility on such information.\footnote{A serious classified information-making one's defenses question would arise: Suppose the court concludes, based upon classified information about the interrogation process, that that process is well designed to ensure sufficient reliability. Suppose also, however, that the court denies access by the defendant to this classified information and the defendant is therefore unable to respond to and argue against that conclusion. Has not the defendant been denied the right to make his defenses?}

The arguments advanced in Moussaoui and Paracha for the circumstantial reliability of the first hearsay level of interrogation statements—whether grounded in intrinsic or extrinsic grounds—do not appear to be strong enough to warrant admissibility under the Federal Rules, neither under a specific exception nor under the residual exception of Rule 807. Claims that the purpose of the interrogations is to obtain intelligence and that the questioners’ motivation, expertise and methods enable them to obtain truthful information and gauge when it is truthful—though interesting and creative—do not in the end seem to be logically strong enough or to be supported by sufficient factual information to overcome the “presumptive unreliability that attaches to accomplices’ confessions that shift or spread blame . . . when the government is involved in the statements’ production, and when the statements describe past events and have not been subjected to adversarial testing.”\footnote{Lilly v. Virginia, 527 U.S. 116, 137 (1999).}

V. CONCLUSION

The rationale articulated by the judges in Moussaoui-Paracha provides the government with a line of arguments to
use in federal court terrorism prosecutions when offering into evidence hearsay statements, which are obtained by the government through questioning for intelligence purposes and which incriminate the defendant. More specifically, it suggests a pathway to explore confrontation-testimonial concerns and reliability-hearsay issues under the Federal Rules in regard to statements obtained from subsequently unavailable declarants.

The weakest link in the argument involves the application of the rationale to the admissibility question at the first level of hearsay—that is, the admissibility of the actual statements made by the interrogatee. The arguments that the substitutions are nontestimonial under Crawford and Davis and that all but the first level of hearsay in the reports/summaries/substitutions meet the standards of Rule 803(8), while certainly far from conclusive, have a reasonable chance of being successful.\footnote{But success here does not coincide with admissibility. Under Fed. R. Evid. 803(8), unless the sources have sufficient trustworthiness, the evidence does not come in. See Fed. R. Evid. 803(8).}

While the suggested arguments that the government might make regarding the first level hearsay issue are creative, it seems unlikely that in a normal trial setting the government would (or should) prevail on that issue.\footnote{However, the same type of background concerns that may have influenced the Moussaoui and Paracha courts to articulate and then apply a reliability rationale in a compulsory process context—that is, concerns about releasing dangerous terrorists—are likely also to be present in future terrorism prosecutions. Potentially, these same kinds of concerns may exert some influence on judges, leading them to take advantage of the type of doctrinal lifeline that Moussaoui-Paracha might be seen as providing regarding confrontation/hearsay issues. Accordingly, it would be foolish to make a firm prediction that the government would be unlikely to succeed on these issues in future terrorism trials in the federal courts.} The court’s argument—that the statements made by the interrogatees have circumstantial reliability because the government interrogators had a strong motivation to obtain accurate information or because the process of interrogation was “well designed” to produce accurate information—runs directly counter to traditional judicial concerns about the risks that arise from government interrogations, as reflected both in Justice Scalia’s opinion in Crawford as well as Professor Berger’s views in her 1992 law review article.

The Moussaoui reliability assertions were made in a compulsory process setting. There may be doubts whether even in that setting it should be permissible to attribute reliability to statements that do not have sufficient indicia of
circumstantial reliability. Be that as it may, we should not be misled by reliability attributions made in a compulsory process setting to conclude that such statements should necessarily be admissible under hearsay rules against a criminal defendant.\textsuperscript{75}

It is noteworthy that, in the end, the hearsay statements at issue are most likely to founder not on confrontation grounds as reflected in \textit{Crawford} and \textit{Davis}, but rather on old-fashioned reliability concerns, as now only reflected in federal and state rules of evidence. The message is that while \textit{Crawford} and \textit{Davis} set up an additional barrier to admissibility framed in terms of the testimonial-nontestimonial distinction, a reliability standard also continues to be applicable, based not in the Constitution but on the rules of evidence.\textsuperscript{76}

\begin{footnotesize}
\begin{footnotes}{\footnotesize
\footnotes{\textsuperscript{75} See Peter Westen, \textit{Confrontation and Compulsory Process: A Unified theory of Evidence for Criminal Cases}, 91 HARK. L. REV. 567, 627 n.167 (1978) (concluding that in determining whether exculpatory evidence has sufficient assurances of reliability to be admissible under compulsory process doctrine one should refer as a benchmark to whether incriminating evidence "would be deemed to possess sufficient 'indicia of reliability' to be admissible against the defendant under the due process clause" (citations omitted)).

Professor Westen thus reasons from the reliability of incriminating evidence to the reliability of compulsory process-exculpatory admissible evidence. The issue posed by our hypothetical situation is the reverse: the question is whether sufficient reliability for admissibility of compulsory process exculpatory evidence should therefore be deemed sufficient reliability to make the evidence admissible where it is incriminatory of the defendant and raises confrontation-hearsay issues. At other places in the article, Professor Westen seems to indicate that the standards of admissibility under both the Confrontation and Compulsory Process Clauses should be the same. See \textit{id.} at 601 ("[T]he principles of confrontation and compulsory process are substantially identical . . . . While the prosecution is under a further obligation to present its evidence in reliable form, it is compelled to do so not by the confrontation clause but by the due process clause."); see also \textit{id.} at 598 ("The due process clause prohibits the state . . . from using any single item of evidence against a defendant which is inherently too unreliable for rational evaluation by the jury.").

}
\end{footnotes}
\end{footnotesize}
Scientific Evidence as Foreign Law

Edward K. Cheng†

I. INTRODUCTION

In his dissent in Daubert v. Merrell Dow Pharmaceuticals, Chief Justice Rehnquist appeared skeptical as to whether federal judges could fulfill the gatekeeping role that the majority had constructed in its watershed opinion.¹ In a sense, Margaret Berger’s defining contribution to the evidence world since Daubert has been to prove Chief Justice Rehnquist wrong. Her unflagging commitment to promoting science education for judges, whether through her Science for Judges program, her work on the Reference Manual for Scientific Evidence, or her work with the National Academy of Sciences, has set a shining example of how academic efforts can help solve, or at least positively impact, real world problems.

The occasion to write for this festschrift in celebration of my colleague, friend, and mentor thus seemed to cry out for a contribution that encompassed both of these attributes: a topic that involved science and judging, as well as one that held academic interest yet had practical implications. I hope that the following succeeds in this regard, but even if it does, I can take only partial credit, for its success would be through following Margaret’s example.

Most contemporary debates about scientific evidence focus on admissibility under Daubert and the Federal Rules of Evidence. That bias is quite understandable—after all, it is the framework imposed by the United States Supreme Court. Daubert, however, rests on a fundamental assumption: that courts should treat scientific facts like any other adjudicative facts ultimately left to the jury. Perhaps the involvement of

† Professor of Law, Brooklyn Law School. Thanks to Elissa Berger and Robin Effron for helpful suggestions, and to LaToya Best for diligent research assistance. Generous support was provided by the Brooklyn Dean’s Summer Research Fund.

¹ Daubert v. Merrell Dow Pharm., 509 U.S. 579, 600 (1993) (Rehnquist, C.J., dissenting) (“I defer to no one in my confidence in federal judges; but I am at a loss to know what is meant when it is said that the scientific status of a theory depends on its ‘falsifiability,’ and I suspect some of them will be, too.”).
specialized knowledge requires judges to act as gatekeepers to ensure some basic level of reliability, but under Daubert, scientific facts are still just facts.

As I will argue, scientific facts fit awkwardly into the conventional framework for conceptualizing and regulating the proof of adjudicative facts. For one thing, scientific facts are rarely ever unique to the case at hand. They are instead often applicable to a variety of cases, and thus ideally should be decided uniformly. At the same time, proof of scientific facts generally depends on an entire body of knowledge, rather than a specific witness or piece of physical evidence. These attributes as well as others suggest that we should think carefully about the framework for scientific factfinding.

Consequently, in this contribution, I look not at how Daubert does or should operate, but rather how the legal system should treat scientific facts more fundamentally. In particular, I suggest that proving scientific facts has much in common with proving foreign law. This perspective shift could prove fruitful for understanding and addressing many of the problems in scientific evidence today. More importantly, the procedural mechanisms developed by conflicts-of-law scholars to handle proof of foreign law can be adapted to the scientific evidence context.

II. FACTUAL FRAMEWORKS

A. The Law-Fact Distinction

Lawyers are intimately familiar with the law-fact distinction and the many implications the dichotomy entails. Perhaps law and fact are not theoretically distinct, but in practice, the distinction makes all the difference. Consider the implications of labeling an issue as a fact question. In an ordinary trial, the jury is the decisionmaker for facts, and the rules of evidence govern the process of proof. Adversarial

---

2 See Ronald J. Allen & Michael S. Pardo, The Myth of the Law-Fact Distinction, 97 Nw. U. L. Rev. 1769, 1769-70 (2003) (arguing that laws and facts have no “essential difference,” and that the distinction is rather a function of pragmatic considerations such as the identity of the factfinder and whether the fact has general or specific import).

3 Conventionally, the rules of evidence govern only jury trials, with judges in bench trials empowered to give erstwhile inadmissible evidence whatever weight they feel fit. For an insightful article about why judges may also benefit from a rule-based evidentiary framework, see Frederick Schauer, On the Supposed Jury-Dependence of Evidence Law, 155 U. Pa. L. Rev. 165 (2006).
system values are also in full swing, with both the judge and the jury assuming a largely passive role. Factfinders ideally have no preexisting knowledge of any litigation-specific facts, and independent research is strictly prohibited. After trial, reviewing courts treat factual findings with the highest deference, but although such findings are binding on the parties through res judicata, there is no stare decisis per se—later parties in other trials are free to relitigate the issues.

The process of finding law operates in sharp contrast. Judges determine the law. They are supposed to know the law, and in many instances, judges independently research relevant law and legal theory, unencumbered by any rules of proof. After trial, appellate courts treat lower court legal determinations with no deference at all. However, stare decisis will bind future parties to the legal decisions made in the present case, in part because law is neither party-specific nor “owned” by the parties, but rather is part of a broader scheme of justice with implications that go beyond the present case.

B. Tensions

Despite its popularity and practical usefulness, the law-fact distinction is hardly a clean one. Over the years, various

---


5 Traditional rules requiring mutuality for issue preclusion to operate have of course been abandoned. 18A CHARLES ALAN WRIGHT, ET AL., FEDERAL PRACTICE AND PROCEDURE § 4464 (2d ed. 2009). However, imposition of “nonmutual preclusion is . . . allowed only if ‘the party against whom an estoppel is asserted had a full and fair opportunity to litigate’ in the first action,” id. (quoting Blonder-Tongue Labs. v. Univ. of Ill. Found., 402 U.S. 313 (1971)), and offensive nonmutual preclusion is permitted only if the trial court determines such imposition to be fair, id. (citing Parklane Hosiery Co. v. Shore, 439 U.S. 322 (1979)).

6 As the Latin phrase goes, jura novit curia, or “the court knows the law.”

7 See generally Amanda Frost, The Limits of Advocacy, 59 DUKE L.J. 447, 461-67 (2009) (discussing, among other things, the ability of courts to raise jurisdictional issues or “extraordinary” merits issues sua sponte).

8 See U.S. Bancorp Mortgage Co. v. Bonner Mall P’ship, 513 U.S. 18, 26-27 (1994) (“Judicial precedents are presumptively correct and valuable to the legal community as a whole. They are not merely the property of private litigants and should stand unless a court concludes that the public interest would be served by a vacatur.” (quoting Izumi Seimitsu Kogyo Kabushiki Kaisha v. U.S. Philips Corp., 510 U.S. 27, 40 (1993) (Stevens, J., dissenting))).
“hybrid” issues have surfaced that fit uncomfortably into the dichotomy and expose its tensions. For example, as Kenneth Culp Davis noted long ago, the category of facts subdivides between adjudicative facts, which are facts in the traditional sense, and legislative facts, which courts use to interpret or develop the law. While adjudicative facts may be appropriately subject to the usual strictures, legislative facts, being part of legal inquiry and the judicial realm, should and do face fewer restrictions. Indeed, building on Davis’s distinction, John Monahan and Laurens Walker have argued that social science research, a species of legislative fact, should be treated akin to “legal precedent under the common law.”

Another orphan of the law-fact regime is foreign law. Here, I am not referencing the contemporary constitutional controversy, which asks whether American courts may legitimately use foreign laws as persuasive or moral authority in interpreting the Constitution. Instead, the relevant foreign law problem for our purposes is the more pedestrian one of how to prove the content of foreign law. In today’s globalized world, courts commonly encounter cases that are governed by the laws of another jurisdiction. The problem becomes how courts determine what French law or Chinese law says about the matter at hand.

11 Compare, e.g., Roper v. Simmons, 543 U.S. 551, 575-79 (2005) (discussing global trends in capital punishment for juvenile offenders), and id. at 604-05 (O’Connor, J., dissenting) (arguing that “the Court has consistently referred to foreign and international law as relevant to its assessment of evolving standards of decency”), with id. at 624 (Scalia, J., dissenting) (rejecting the idea of using foreign law). A sizable literature on the issue of foreign and international law in constitutional jurisprudence—too voluminous to catalog here—has developed in the wake of Roper and other cases.
12 See generally Aurora Bewicke, The Court’s Duty to Conduct Independent Research into Chinese Law: A Look at Federal Rule of Civil Procedure 44.1 and Beyond, 1 CHINESE L. & POL’Y REV. 97 (2005) (reporting recent cases involving the application of foreign law). Some courts evade the foreign law problem altogether by simply applying the law of the forum, a practice that commentators have criticized. See, e.g., Vishipco Line v. Chase Manhattan Bank, N.A., 660 F.2d 854, 860 (2d Cir. 1981) (suggesting that applying the law of the forum is permissible, even if technically incorrect under choice-of-law rules, if neither party objects); RESTATEMENT (SECOND) CONFLICTS OF LAW § 136 cmt. h (1971) (“When both parties have failed to prove the foreign law, the forum may say that the parties have acquiesced in the application of the local law of the forum.”); Roger J. Miner, The Reception of Foreign Law in the U.S. Federal Courts, 43 AM. J. COMP. L. 581, 583 (1995) (criticizing the practice of ignoring foreign law and simply applying the convenient law of the forum).
One may be tempted to treat foreign law no differently than the law of the forum—after all, law is law. But a moment's reflection reveals the problem to be trickier than it would seem at first glance. Unlike in the case of domestic law, the judge in a foreign law case does not bring a lifetime of experience and expertise to the task.\textsuperscript{13} The relevant statutory and case materials are likely in another language and out of an entirely different legal tradition.\textsuperscript{14} The judge will therefore need the help of some type of factfinding process, most often through an expert provided by the parties or appointed by the court.

Questions about scientific facts present similar problems.\textsuperscript{15} Although they are treated as facts, general scientific facts, such as whether a scientific method like DNA typing is valid, or whether a substance causes cancer, fit poorly into the ordinary factual framework. The proof process for facts is built largely around the assumption that fact determinations are specific to the case. Usually, this perspective makes sense, because adjudicative facts are of little interest beyond the litigants at bar. No other institution will ever have more information or be better equipped than the jury to decide a factual issue. Second-guessing the jury only creates inefficiency.

Scientific facts, however, are different. Being general truths, they ideally should apply consistently from one case to another.\textsuperscript{16} In addition, scientific facts are easily subject to external scrutiny. Unlike ordinary facts, in which no one is the wiser, with scientific facts, whole communities of scientists stand ready to challenge erroneous court findings. The legal system therefore has an important broader interest in establishing scientific facts accurately beyond doing justice in the individual case.

\textsuperscript{13} See Gregory S. Alexander, The Application and Avoidance of Foreign Law in the Law of Conflicts, 70 NW. U. L. REV. 602, 603-04, 630-31 (1975) (noting that judges are far more likely to be ignorant of foreign law and lack the context necessary to interpret new provisions).

\textsuperscript{14} See John Henry Merryman & Rogelio Pérez-Perdomo, The Civil Law Tradition (1985) 61-67 (describing the idea of “legal science” in German jurisprudence and how it sharply diverges from American legal realism).

\textsuperscript{15} To be perfectly precise, the term “scientific facts” here does not include individual applications of science, which are case-specific and far more like ordinary adjudicative facts. See, e.g., In re UNISYS Sav. Plan Litig., 173 F.3d 145, 161 (3d Cir. 1999) (Becker, C.J., dissenting) (drawing a distinction between general methodology questions which are for the judge, and questions about the reliability of a specific expert witness, which are for the jury).

\textsuperscript{16} See supra note 15.
C. Finding a Home for Scientific Evidence

Given the problems with the factual regime, one option is to treat scientific evidence as akin to law, along the lines of Monahan and Walker.\textsuperscript{17} Here, however, the fit remains problematic. The proof process for law is relaxed precisely because judges are assumed to be well-versed and experienced in interpreting the laws of their home jurisdiction. Yet, judges rarely if ever will have such a high comfort level with scientific issues.\textsuperscript{18}

But if scientific evidence is neither fish nor fowl, how should we handle it? The standard response for a law review article is to propose some kind of third, hybrid category. The proposal in turn prompts the reader to roll her eyes, and rightly so—after all, the whole point of a dichotomy is to make rough divisions, and adding a third or fourth classification is rarely justified. But in this case, there actually already exists such a hybrid category, and it is the one for foreign law.

Foreign law and scientific facts are neither law nor fact for roughly the same reasons. Approaching them as factual questions neglects their status as generalized inquiries subject to external verification. Approaching them as legal questions ignores judges’ profound lack of expertise and experience in the substantive areas. Both inquiries thus fall into a no-man’s land between law and fact. Indeed, they are so similar that the governing doctrines should arguably cohere.

The good news is that conflicts-of-law scholars have grappled with the problem of proving foreign law for some time. Indeed, the poor fit between the law-fact dichotomy and foreign law questions has borne itself out in actual doctrinal wrangling. At common law and historically in the federal courts, questions of foreign law were treated as questions of fact so that pleading requirements, the rules of evidence, and

\textsuperscript{17} Monahan & Walker, supra note 10, at 488.

\textsuperscript{18} For example, consider Judge Kozinski’s reaction to the Supreme Court’s Daubert decision:

\textit{As we read the Supreme Court’s teaching in Daubert, therefore, 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 determine whether those experts’ proposed testimony amounts to “scientific knowledge,” constitutes “good science,” and was “derived by the scientific method.” . . . Mindful of our position in the hierarchy of the federal judiciary, we take a deep breath and proceed with this heady task.}

\textit{Daubert v. Merrell Dow Pharms., 43 F.3d 1311, 1316 (9th Cir. 1995).}
adversarial values all applied.” Yet, in 1966, Federal Rule of Civil Procedure 44.1 did the unthinkable—it recharacterized foreign law as a question of law, opening the door to de novo appellate review and independent judicial investigations.\textsuperscript{20}

Courts have arguably never fully adopted the fact or law framework exclusively. Instead, the regime governing foreign law questions often ends up as a hybrid, mixing and matching procedures and requirements from both categories of proof. For example, despite defining foreign law to be a question of law, federal courts have effectively held that a failure to provide sufficient evidence of foreign law remains a valid ground for dismissal.\textsuperscript{21} Along similar lines, while federal appellate courts have outwardly encouraged judges to do independent research on foreign law,\textsuperscript{22} in practice, judges remain reluctant,\textsuperscript{23} and perhaps more tellingly, rely substantially on experts, something they would almost never do for domestic law.

Such hybridization also appears outside the federal context. Texas has “a hybrid rule by which the presentation of the foreign law to the court resembles the presentment of evidence but which ultimately is decided as a question of law.”\textsuperscript{24} Other states handle foreign law questions primarily through judicial notice,\textsuperscript{25} but as one leading casebook observes, for cases involving “foreign legal system[s] . . . alien in language and structure,” courts will often rely heavily on party presentation and decline to do independent research (evoking “fact”), yet


\textsuperscript{20} \textit{Fed. R. Civ. P. 44.1 & advisory committee’s notes; see also Tex. R. Evid. 203. But see Griffin v. Mark Travel Corp., 724 N.W.2d 900, 902 (Wis. Ct. App. 2006) (noting that Wisconsin maintains the common law classification of foreign law as fact); Amsellem v. Amsellem, 730 N.Y.S.2d 212, 215 (N.Y. Sup. Ct. 2001) (holding that “the interpretation of French law is an issue of fact that can be resolved at trial”).

\textsuperscript{21} See Esso Standard Oil S.A. v. S.S. Gasbras Sul, 387 F.2d 573, 581 (2d Cir. 1967) (holding that plaintiff “failed in its burden of proof” regarding foreign law and that even under Rule 44.1, plaintiff failed to show “it ha[d] a good cause of action”), cited in Adams v. Arabian Am. Oil Co., No. 92-35028 1993 U.S. App. LEXIS 25448, at *7-8 (9th Cir. Sept. 28, 1993) (unpublished opinion) (noting that although “there is no ‘burden of proof’ in the evidentiary sense with respect to foreign law . . . the plaintiff who pleads foreign law still must successfully persuade the court that he has a good cause of action”).

\textsuperscript{22} See Twohy v. First Nat’l Bank of Chi., 758 F.2d 1185, 1193 (7th Cir. 1985).

\textsuperscript{23} \textbf{RUDOLF B. SCHLESINGER, ET AL., COMPARATIVE LAW} 73 (6th ed. 1998).

\textsuperscript{24} See, e.g., Long Distance Int’l, Inc. v. Telefones de Mex., 49 S.W.3d 347, 351 (Tex. 2001) (discussing \textit{Tex. R. Evid. 203}).

\textsuperscript{25} \textit{E.g., N.Y. C.P.L.R. 4511 (2007).}
dispense with the ordinarily restrictive rules of evidence (evoking “law”).

The regimes for handling foreign law are by no means uniform: the particular breed of hybridization differs from one jurisdiction to another. Considered together, however, they offer a new perspective on the proof of scientific evidence, and implicitly suggest some avenues for reform. The next Part draws out some of these implications.

III. APPLYING THE FOREIGN LAW MODEL

What would be the ramifications of imposing a foreign law model on scientific evidence? To get a sense, this Part maps some of the more fundamental features of a typical foreign law framework onto the scientific evidence context. In the abstract, some of the resulting proof requirements and procedures may appear a bit radical, but a closer look shows them to be sensible, and in some cases, not all that different from judges’ natural inclinations.

A. Proof at Trial

A hybrid treatment of scientific facts would undoubtedly result in important changes to the process of proof at trial. For example, one of the key changes that arose when foreign law shifted from a fact to a law regime was the relaxation of the focus on oral presentation of evidence. The move to a law regime also affected the applicability of the rules of evidence, judicial notice, and pleading requirements.

1. Oral Presentation

Whatever can be said about the merits of conventional direct and cross examination, they almost certainly do not

---

26 Schlesinger, supra note 23, at 69.
27 Id. at 95-98 (surveying the states and concluding that a plurality follow FED. R. CIV. P. 44.1, but that a significant number adopt some type of judicial notice approach, and a small minority continue to treat foreign law as fact).
28 E.g., FED R. CIV. P. 44.1 (“In determining foreign law, the court may consider any relevant material or source, including testimony, whether or not submitted by a party or admissible under the Federal Rules of Evidence.”); TEX. R. EVID. 203 (“The court, in determining the law of a foreign nation, may consider any material or source, whether or not submitted by a party or admissible under the rules of evidence, including but not limited to affidavits, testimony, briefs, and treatises.”).
apply to scientific questions. Consider how good instructors and researchers generally convey new scientific ideas. Ideas are developed through a (hopefully) carefully constructed lecture in which background concepts build up to a main thesis. And as any veteran of a science course knows, sound understanding of that lecture almost always requires lengthy and prior preparation with reading materials. This recipe undoubtedly continues to hold true in the courtroom, where cases often turn on a stack of scientific articles (or lack thereof). The choppy back-and-forth of ordinary courtroom testimony is suboptimal and confusing, and knowledgeable experts unfamiliar with legal examination methods become easily derailed by the attorneys.

2. Rules of Evidence

The applicability of the rules of evidence also changes considerably for the better. One of the biggest weapons in the rules of evidence is the hearsay rule, which requires that witnesses testify only about matters for which they have personal knowledge. Scientific witnesses will almost never satisfy the rule, since science is a collaborative process, taking place over considerable distances and time. Expert testimony rules, of course, create sizable exceptions to the hearsay rule, allowing witnesses to use inadmissible evidence for the purpose of reaching their conclusions, but the strong presumption is

29 See Harold L. Korn, Law, Fact, and Science in the Courts, 66 COLUM. L. REV. 1080, 1086 (1966) (“The larger issue, beyond criticism of obstacles to expert testimony, is whether oral communication is at all conducive to correct determination of complicated scientific and technological issues.”).

30 Indeed, the use of mini-lectures to juries has become increasingly common in scientific cases. See, e.g., Marvin J. Garbis, Aussie Inspired Musings on Technological Issues—Of Kangaroo Courts, Tutorials & Hot Tub Cross-Examination, 6 GREEN BAG 141, 144 & n.16 (2003) (recounting a federal patent trial in which the judge had considered having a “tutorial expert . . . give[] the jury an introductory tutorial lecture”).

31 See, e.g., State v. Erickson, 574 P.2d 1, 6 (Alaska 1978) (questioning, in a case involving cocaine’s proper classification, whether a hearing would provide better evidence than looking at briefs and judicial research); Neil Vidmar & Shari Seidman Diamond, Juries and Expert Evidence, 66 BROOK. L. REV. 1121, 1152-53 (2001) (discussing psychological studies showing that jurors who receive summaries of expert testimony “were more likely to make clear distinctions . . . and recall more trial-relevant information”).

32 FED. R. EVID. 703 (“If of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject, the facts or data need not be admissible in evidence in order for the opinion or inference to be admitted.”).
that these bases are never revealed to the factfinder.\textsuperscript{33} Even under the learned treatises exception to the hearsay rule, scientific journal articles “may [only] be read into evidence but may not be received as exhibits.”\textsuperscript{34} The end result is that jurors are structurally prevented from grappling with the basis of an expert’s opinion, forcing them to either believe the expert or not. This “deference” model of expert testimony is rife with danger, particularly since relying on traditional cues for assessing witness credibility is not necessarily a sound method for assessing scientific experts.\textsuperscript{35} Furthermore, there is almost no reason to judge scientific facts upon the idiosyncrasies of a particular scientist. Unlike with traditional facts, in which eyewitnesses are not generally fungible, with scientific facts the witnesses typically are. The expert is a synthesizer of vast quantities of information, and it is the information and not the expert that ideally should be the focus of the inquiry.

3. Judicial Notice

For similar reasons, judicial notice practice also improves. For adjudicative facts, judicial notice involves an extremely high bar. Words like “indisputable,” or “beyond controversy” litter the landscape, discouraging parties and courts from deviating from the ordinary presentation of

\textsuperscript{33} FED. R. EVID. 703 (“Facts or data that are otherwise inadmissible shall not be disclosed to the jury . . . unless the court determines that their probative value in assisting the jury to evaluate the expert’s opinion substantially outweighs their prejudicial effect.”).

\textsuperscript{34} FED. R. EVID. 803(18). The federal version of the learned treatises exception is in fact liberal when compared to some state schemes. See, e.g., CAL. EVID. § 1341 (creating hearsay exception only for “books of science . . . made by persons indifferent between the parties . . . when offered to prove facts of general notoriety and interest”); Glenn Koppel, Re: Scientific Evidence as Hearsay, Evidence Listserv Discussion, Feb. 24, 2010 (on file with author) (discussing the narrowness of the California exception).


\textsuperscript{36} But see Sanja Kutnjak Ivkovi & Valerie P. Hans, Jurors’ Evaluations of Expert Testimony: Judging the Messenger and the Message, 28 LAW & SOC. INQUIRY 441 (2003) (presenting research showing that jurors “consider both the messenger and the message in the course of evaluating the expert’s credibility”); Daniel W. Schuman & Anthony Champagne, Removing the People From the Legal Process: The Rhetoric and Research on Judicial Selection and Juries, 3 PSYCHOL. PUB. POL’Y & L. 242, 253-54 (1997) (reporting research suggesting that jurors evaluate experts based on “a very sensible set of considerations—the expert’s qualifications, reasoning, factual familiarity, and impartiality,” and that they “attempt to go beyond superficial considerations”).
evidence. For questions of foreign law, however, the strictures become considerably more relaxed, a state-of-affairs that arguably better suits scientific evidence. Since scientific facts are generalized facts, the court can afford to be more inquisitive, and judicial notice need not demand iron-clad evidence.

4. Pleading

Finally, the hybrid model for foreign law retains (at least in spirit) the pleading requirements from the fact regime. As previously noted, failure to provide evidence of foreign law therefore can provide grounds for dismissal. Retaining these pleading requirements in the scientific context seems similarly sensible. When dealing with the law of the forum, extensive pleading is not required because everyone is presumed to have access to the governing law, and because the court has both expertise and comprehensive research mechanisms at its disposal. This is of course not true for scientific facts. Placing burdens on the moving party thus seems eminently reasonable.

B. Seeking Information Beyond the Parties

The aforementioned changes to the proof process may be the more fundamental implications of imposing a foreign law model on scientific facts, but perhaps the more exciting ones are the inquisitorial mechanisms that the foreign law model suggests. In ascertaining the foreign law applicable to a case, some courts have demonstrated the power of independent judicial research, court-appointed experts, and external institutions. To be sure, some judges still resist using these mechanisms because they run sharply against adversarial

---

37 FED. R. EVID. 201 (“A judicially noticed fact must be one not subject to reasonable dispute . . . .”); FED. R. EVID. 201 advisory committee’s notes (noting that for adjudicative facts, “[a] high degree of indisputability is the essential prerequisite”).

38 My embrace of a conventional pleading regime for scientific facts may initially appear in tension with Margaret Berger’s well-known proposal to dispense with proof of general causation in toxic tort cases. See Margaret A. Berger & Aaron D. Twerski, Uncertainty and Informed Choice: Unmasking Daubert, 104 MICH. L. REV. 257 (2005); Margaret A. Berger, Eliminating General Causation: Notes Towards a New Theory of Justice and Toxic Torts, 97 COLUM. L. REV. 2117 (1997). Margaret’s thesis, however, is a substantive one—namely that tort law should rethink its focus on causation and refocus on things like culpability and failures to test. My focus, in contrast, is on the process of proof presuming no change in the substantive law.
norms, but their explicit recognition in the foreign law context has promoted greater acceptance, a process that could improve how courts handle scientific evidence.

1. Court-Appointed Experts

Court-appointed experts are perhaps the most modest suggestion to come from the foreign law hybrid model. The Federal Rules of Evidence explicitly discuss and allow the use of court-appointed experts in scientific evidence cases.39 Some courts have used them over the years,40 and the Federal Judicial Center has published material describing, facilitating, and encouraging their use.41 Indeed, even Justice Breyer, writing in Joiner v. General Electric Co. managed to make a pitch.42

That said, the reality on the ground is that court-appointed experts are rarely used.43 For some judges, the idea of a neutral expert is anathema, whether because it is inconsistent with the adversarial process, or because it smacks too much of judicial abdication. But for most, the difficulties of finding, funding, and accommodating a court-appointed expert are simply not worth the perceived benefits, so trudge on the judge (or jury) must. An added endorsement from the hybrid model, however, may be a welcome boost of legitimacy, encouraging judges to do more along these lines.


40 See, e.g., Debra L. Worthington et al., Hindsight Bias, Daubert, and the Silicone Breast Implant Litigation, 8 PSYCH. PUB. POL’Y & L. 154, 162 (2002) (listing several high-profile examples, including the breast implant, DES, asbestos, and Parlodel litigations).


2. Independent Judicial Research

The next level of reform deals with independent judicial investigations. Beyond considering party-provided information and arguments, judges handling questions of law may generally conduct independent research. As previously noted, a motivation for this exception to the usual rule against ex parte behavior is that courts should correctly apply legal rules regardless of the parties' positions, especially since those decisions will become precedents in future cases. Even for core legal questions this practice has its detractors, but it is reasonably well accepted.

Independent judicial research can be quite useful for judges attempting to understand scientific testimony as well. Party testimony in this sphere tends to become a battle of the experts with the court placed in the unenviable position of mediating between two or more well-credentialed scientists. In this context, the natural inclination of any beleaguered decisionmaker is to do independent library research on his own. I have argued elsewhere that this practice is both legitimate and desirable, despite the obvious sacrifice to adversarial norms. The judges themselves, however, are evenly split, although moving to a hybrid model may legitimate the practice as it has in the foreign law context.

3. External Institutions

Perhaps the most radical reform suggested by the foreign law model is the use of an external institution with greater expertise for determining scientific facts. In the foreign law context, this general idea surfaces in the use of comparative law research centers and certification procedures. For example, in Germany and France, courts have historically used comparative law centers, notably the Max Planck Institute for Foreign and International Private Law and the

---

45 The classic articulation of this problem is found in Learned Hand, Historical and Practical Considerations Regarding Expert Testimony, 15 HARV. L. REV. 40 (1901). As Judge Hand asks, “how can the jury judge between two statements each founded upon an experience confessedly foreign in kind to their own? It is just because they are incompetent for such a task that the expert is necessary at all.” Id. at 54.
47 Id. at 1276-77 & figs.1 & 2.
French Center of Comparative Law, to gather information on “unfamiliar foreign laws.” Somewhat analogously, federal courts, when confronted with ambiguous substantive issues of state law under diversity, often certify such questions to state supreme courts.

Concededly, no such procedures appear to exist in American courts for gathering information about foreign law. Use of comparative law centers is a European phenomenon, whereas certification surfaces only with regard to state law issues under *Erie.* Yet, on a conceptual level, the procedures seem eminently sensible. If a court faces a difficult or ambiguous question of foreign law and is at a loss as to how to resolve it, whom better to ask than a foreign court?

The analog to these mechanisms in the scientific evidence context is to certify questions to relevant scientific bodies asking for advice. Clearly the context is not exactly the same, since a foreign or state supreme court has conclusive authority to declare the rule for its jurisdiction, whereas scientific bodies obviously do not. Nonetheless, an opinion on the state of science from, for example, the National Research Council of the National Academy of Sciences, can be extremely helpful, if not de facto conclusive in resolving scientific questions in court.

One charge against such certification is its lack of democratic accountability, and that may indeed be a serious cost of the practice. However, as a practical matter, previous reports by the National Academy of Sciences on scientific matters with legal import have been well received and influential in legal circles. The most famous instance is

---


50 Keller, *supra* note 49, at 183 & n.253 (reporting that there is currently “no procedure by which federal courts can certify a difficult foreign legal question to the relevant foreign court,” except for Puerto Rico, which is “clearly a unique situation”).

51 *Id.* at 184-85 (proposing the application of *Erie*-type procedures to the foreign law context).

52 See generally D.H. Kaye, *The NRC Bullet-Lead Report: Should Science Committee Make Legal Findings?,* 46 Jurimetrics J. 91, 104-05 (2005) (raising the question whether science committees like the National Academies should make legal determinations in addition to scientific findings).
perhaps the DNA study,53 but more recent panels on lead bullet analysis54 and the like have been similarly successful.55 All of these previous reports have arisen independent of any formal judicial request or funding, but one wonders if a more formal link could prove beneficial for both sides.56

C. Decisionmakers

The last area in which the foreign law model might offer suggestions for scientific evidence inquiries is in the decisionmaking process itself. Under a foreign-law-as-fact framework, the jury was the finder of foreign law, and appellate courts reviewed these decisions with a high level of deference. With the shift to a more law-oriented framework, foreign law became the province of the judge and was reviewed de novo on appeal.

1. Factfinder

Scientific evidence as it currently stands already occupies the middle ground in terms of decisionmaker. Nominally, scientific facts are facts for the jury. Daubert is merely a reliability screen, no different than other admissibility inquiries under the rules of evidence. But in practice, as everyone knows, Daubert has had monumental significance on the way litigants prove scientific facts in court. After all, excluding an opponent’s scientific expert effectively negates his ability to present any scientific evidence at all, and thus cases often live and die at Daubert hearings.57

55 Perhaps the most ambitious project of all has been the recent National Research Council report on forensic science, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009). Its impact still remains to be seen.
56 My conversations with Margaret, who has of course served on some of these efforts, suggests that they are exceptionally expensive and impractical for all but the most controversial and pressing problems. One wonders, however, if perhaps less expensive versions could be similarly organized.
Adopting the foreign law model may therefore involve a change in factfinder that is more significant from a formal standpoint than a practical one. For all intents and purposes, judges have already claimed a lion’s share of the scientific factfinding process. Undoubtedly, however, adopting the foreign law model would complete the transformation and give it greater transparency.

2. Appellate Review

At present, federal appellate courts review scientific reliability determinations under *Daubert* only for abuse of discretion. The application of an abuse-of-discretion standard is perfectly in line with appellate review standards for other evidentiary rulings, but critically misses the generality that distinguishes scientific from ordinary adjudicative facts. The current *Joiner* doctrine contemplates having one case find a scientific or forensic method sufficiently reliable to be admissible, while a second case does not. The problem is that there can be only one right answer, and precedent should reflect that.

Again, shifting to a foreign law model would help legitimate a change to de novo review, which would have appellate courts perform their established role in ensuring uniformity and consistency among lower courts. Deferential review is well-established for facts and the evidentiary rules that govern them. By converting the proof of scientific facts into a more law-like hybrid, the benefits of de novo review become more obvious, increasing the feasibility of such a shift.

IV. Conclusion

The foreign law model offers a third option beyond the conventional law-fact distinction for handling the proof of scientific facts. The problems of proving foreign law and proving scientific facts are sufficiently close that the relatively modern scientific evidence field can take advantage of the

59 My co-authors and I argue precisely along these lines in our treatise, and as it happens, many courts have seemingly distinguished case-specific scientific evidence, which should be reviewed deferentially, from “trans-case scientific issues,” which should be reviewed de novo. *Joiner*, however, draws no such distinction, making it problematic and in need of additional sharpening. DAVID L. FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE § 1:34, at 100-03 & n.17 (2009).
wisdom developed over the years by conflicts scholars. In its strong form, the foreign law model provides a specific, concrete model for reform in scientific evidence. But even in its weak form, the linking of foreign law and scientific evidence is a useful thought experiment that provides a launching point for discussion on how we might rethink the process of proof in scientific cases.\(^{60}\)

A final matter worth acknowledging is the obstacles a hybrid vision for scientific evidence might encounter. Doctrinally, the most serious barriers are probably the constitutional ones. The Seventh Amendment guarantee of jury trials, for example, poses problems for the move to a judicial factfinder. One response might be to revisit the arguments in support of Federal Rule of Civil Procedure 44.1's constitutionality,\(^{61}\) but those arguments are far less persuasive in the scientific evidence context, if for no other reason than that scientific facts are still “facts” by any common understanding, and no amount of doublespeak will turn them into “law.” Consequently, any practical reform may need to perpetuate a *Daubert* structure in which the jury remains the nominal factfinder, even if it features high levels of judicial supervision.

Along more cultural lines, as Sam Gross and others have argued, given our adversarial traditions, the legal system often resists and ignores inquisitorial reforms.\(^{62}\) Advocating for each of these reforms in isolation may therefore be too much to ask. However, one way to combat cultural resistance is to start with practices with which judges are familiar and comfortable, and then expand them gradually. For example, *Daubert* itself is somewhat inquisitorial in flavor, since it sharply chastens the conventional, adversarial presentation of expert evidence by imposing a judicial gatekeeper. *Daubert* has become widely adopted beyond the federal system and has unquestionably

---

\(^{60}\) This more chastened view may be the more realistic one, since the problem of proving foreign law is not without continuing difficulties. See Alexander, *supra* note 13, at 630 (suggested that both common law and civil law methods of proving foreign law have been similarly “ineffective”).

\(^{61}\) Miller, *supra* note 19, at 684-88 (discussing various reasons why foreign law, although a question of fact at common law, may not require jury determination under the Seventh Amendment); see also Note, *Proof of the Law of Foreign Countries: Appellate Review and Subsequent Litigation*, 72 HARV. L. REV. 318, 322 n.38 (1958) (“[I]t is arguable that since foreign law was decided by a jury at common law, a federal court is bound by the seventh amendment to give the question to the jury and exercise a narrow scope of review.” (citation omitted)).

influenced the way judges (and evidence scholars) think about scientific evidence. What explains its success? Certainly being a Supreme Court edict helped, but the genius of *Daubert* was framing the reform of scientific evidence as an admissibility problem. Judges are comfortable making admissibility decisions; *Daubert* is only a modest extension.

If judges have (or will) become accustomed to the Rule 44.1 framework for handling foreign law, then shifting scientific evidence questions into this hybrid regime could be more successful than adopting the reforms piecemeal. Lawyers are creatures of analogy, and once comfortable with Rule 44.1, extending it to scientific evidence may only require a minor mental shift.

* * *

Two aspects of Margaret’s scholarship and approach to evidence have always struck me as important guideposts for future generations of evidence scholars. One is aspirational, the other cautionary.

The aspirational thread is that in thinking about evidence and proof, one should avoid tunnel vision and confining oneself exclusively to the “Rules.” No article better demonstrates the rewards of this kind of “outside-the-box” thinking than Margaret’s 1997 article on eliminating general causation.63 Proving causation in toxic tort cases is often a remarkably difficult task from an evidentiary standpoint, resulting in unnecessary social costs.64 Margaret’s solution to the problem, however, is not endless tweaking of the evidentiary doctrines, but rather considering new tort perspectives and changing the substantive tort requirements.65

The cautionary thread is simply that scientific evidence problems are, as Margaret likes to say, “very difficult.” The problems of expert evidence have plagued courts for over two centuries,66 and while as academics we are preternaturally disposed to elegant and grand solutions, those solutions are unlikely to work out. Real life is too messy, and in this business, silver bullets are few and far between. Change and improvement will realistically arise through accretion, not revolution.

---

63 Berger, supra note 38.
64 *Id.* at 2118, 2122-31.
65 *Id.* at 2152.
My hope is that linking scientific evidence with foreign law pays tribute to both these nuggets of wisdom distilled from Margaret's long and distinguished career. On the first score, the link blends evidence law with comparative law, two fields so disparate that I dare say they share few if any scholars in common. On the second, by hewing closely to well-established precedents in the conflicts field and using modest analogies to port them over to scientific evidence, the proposal eschews the grand “solution” in favor of the messier but more gradual and flexible “approach.”

Of course, none of this is to say that Margaret will not have her doubts about my crazy idea, but I would not have it any other way.
Evidentiary Incommensurability

A PRELIMINARY EXPLORATION OF THE PROBLEM OF REASONING FROM GENERAL SCIENTIFIC DATA TO INDIVIDUALIZED LEGAL DECISION-MAKING

David L. Faigman

INTRODUCTION

Scientists typically study variables at the population level, and most of their methodological and statistical tools are designed for this kind of work. The trial process, in contrast, ordinarily concerns whether a particular case is an instance of the general phenomenon. As I have previously observed, “[w]hile science attempts to discover the universals hiding among the particulars, trial courts attempt to discover the particulars hiding among the universals.” This essential difference in perspective between what scientists normally do and what the trial process is ordinarily about has yet to be studied with any degree of rigor—by scientists or lawyers. Yet


2 One exception to this yawning silence is the work of Joseph Sanders, who provides a careful examination of this issue in this volume. See generally Joseph Sanders, Applying Daubert Inconsistently? Proof of Individual Causation in Toxic Tort and Forensic Cases, 75 Brook. L. Rev. 1367 (2010). In addition, the statistical challenges associated with individualizing group data have been examined with considerable sophistication in the context of predictions of violence. See, e.g., Stephen D. Hart et al., Precision of Actuarial Risk Assessment Instruments: Evaluating the ‘Margins of Error’ of Group v. Individual Predictions of Violence, 190 Brit. J. Psychiatry 660 (2007); Douglas Mossman, Analyzing the Performance of Risk Assessment Instruments: A Response to Vrieze and Grove, 32 Law & Hum. Behav. 279,
this phenomenon is endemic to virtually every context in which law and science meet. Indeed, it might be said to be the single greatest obstacle to the law’s rational use of science. The challenges associated with individualizing science, however, are not unique to the law. In fact, in a wide variety of social contexts, empirical research exploring general phenomena are sought to be applied reliably to individual cases. In medicine, for example, research on the effectiveness of various cancer therapies will inform a particular patient’s decision regarding which therapy to choose. In meteorology, research on hurricanes will inform a governor’s decision regarding whether to evacuate a particular city. Indeed, all applied science potentially presents the problem of making decisions about discrete cases based on group data, ranging from aerodynamics to zoology. Different fields have adapted strategies to respond to the evidentiary incommensurability challenge with differing degrees of success. In medical decision-making, for example, evidence-based medicine is one way that doctors have sought to bring data to bear on individual diagnostic and therapeutic judgments. Meteorologists generate computer models that describe the likelihoods associated with a storm’s path and strength. At least from an outsider’s


3 Professor Margaret Berger, to whom this Festschrift and this essay are dedicated, has devoted a substantial portion of her scholarship to navigating the intersection of law and science. The basic incompatibility between much of what science is able to offer and what most courts would like has been an abiding topic in her extraordinary scholarship. See, e.g., Margaret A. Berger, Eliminating General Causation: Notes Towards a New Theory of Justice and Toxic Torts, 97 COLUM. L. REV. 2117, 2129-30 (1997); Margaret A. Berger & Lawrence M. Solan, The Uneasy Relationship Between Science and Law: An Essay and Introduction, 73 BROOK. L. REV. 847, 852-53 (2008).

4 Daniel B. Mark, Decision-Making in Clinical Medicine, in 1 HARRISON’S PRINCIPLES OF INTERNAL MEDICINE 6, 6 (Dennis L. Kasper et al. eds., 16th ed. 2005).

perspective, these efforts have not been so successful that courts would want to borrow them wholesale.\(^6\)

How, and whether, general data can be usefully employed to inform decisions about individual events is a problem that is central to the law’s function. In fact, courts are generally acquainted with the difficulties inherent in employing general scientific data to reach conclusions about specific cases. The primary area in which courts have considered this matter is in medical causation cases where they distinguish routinely between “general causation” and “specific causation.” Courts and legal scholars have not, however, engaged in a careful study of the details and intricacies associated with this matter across the wide spectrum of cases in which it presents itself. In addition, although the courts are passingly familiar with the problem of evidentiary incommensurability, they naturally approach the subject from their own need for information, with little appreciation for how and whether scientists can produce this information. Courts frequently demand empirical answers despite scientists’ inability to provide them.\(^7\) At the same time, scientists involved in the legal process naturally approach the problem of incommensurability from the perspective of their own desire to produce information, with little appreciation for how and whether the courts can effectively use this information.\(^8\) It is

---

\(^6\) For example, Dr. Jerome Groopman cautions against over-reliance on evidence-based medicine, fearing that it “risks having the physician choose passively, solely by the numbers,” rather than rely on the individual circumstances of each patient. JEROME GROOPMAN, HOW DOCTORS THINK 5-6 (2007).

\(^7\) See DAVID L. FAIGMAN ET AL., 3 MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY § 21.6-7, at 27-45 (2008-2009 ed.) (listing cases). Not all science is engaged in describing cause and effect relationships, so “general causation” and “specific causation” are subcategories of what might more properly be labeled “general propositions” and “specific application.” Sometimes general propositions in science will be stated in causative terms, but very often they will be associational, technical, or descriptive. Specific application refers to the determination whether a particular case is an instance, use, or example of general propositions that are supported by research.

\(^8\) Among many possible examples that could be cited, possibly the most obvious is that of predicting violence. Courts call upon experts in myriad contexts to predict future behavior, from probation decisions to capital sentencing, though the best empirical research indicates that such expert opinions remain highly fallible. See John Monahan, A Jurisprudence of Risk Assessment: Forecasting Harm Among Prisoners, Predators, and Patients, 92 VA. L. REV. 391, 405-07 (2006).

\(^9\) Scientists do not generally study how to “individualize” their findings in ways that would make them most helpful for legal usage. This is not meant as a criticism, only an observation. Especially in the social sciences, it is ordinarily sufficient to find a statistically significant effect among college sophomores. Little attention has been paid to how the variables studied might operate in a particular case.
hardly surprising that scientists should study the questions that they are most curious about and able to answer rather than those the law deems most relevant. In short, therefore, the two sides, law and science, perceive incommensurability from their separate vantage points, which largely perpetuates the problem.

This essay jumps into the center of this conundrum. My objective, however, is somewhat unusual. I do not aim to resolve the incommensurability paradox, but rather to ring the fire-bell. Indeed, given the scope and depth of the obstacles presented by evidentiary incommensurability, it is a subject well beyond resolution in the pages provided to me here. My purpose, then, is to explore the paradox in the hope that it will help lay a common framework by which both lawyers and scientists might understand the challenges presented at the intersection of these two great professions. This essay, therefore, contemplates many of the sundry issues that would have to be reckoned with in any subsequent comprehensive effort to bring systematic rationality to the problem of employing group data to decide individual cases. It is divided into two parts. Part I considers scientific hypothesis testing and the inherent population focus of most of that work. While most scientific research focuses on a general population level analysis, results of that work can have very different levels of probative value in regard to informing decision-making at the individual level. Part II examines evidentiary demands in the courtroom and the inherent individualized focus of that process. This part also considers some of the challenges inherent in any attempt to close the evidentiary incommensurability gap between what most science says and what most legal proceedings need to know.

I. HYPOTHESIS TESTING IN SCIENCE

Scientific research is most often conducted from a general and population-based perspective. This is a defining characteristic of the field. However, scientific methods, and the phenomena that scientists study, range widely. Inevitably, the demands of the empirical context dictate which set of research designs are, or might be, available. While studying the effects of depleted biodiversity in the Amazon rainforest and investigating the interaction between neuron and glial cells in a rat’s brain are both scientific endeavors, the methods involved are obviously disparate. Yet, from the law’s
perspective, there may be certain insights that persist across scientific domains in regard to individualizing group data. This section provides a preliminary sketch of the scientific landscape regarding whether certain common denominators might be identified within the process of bringing group data to bear on individual decisions.

The essential question posed in the context of reasoning from the group to the individual is whether a particular case is an instance of the general phenomenon. If smoking causes lung cancer, the individualized query is whether a particular person’s lung cancer was caused by smoking. The degree to which scientific research might be relevant to resolving an individualized fact question varies from complete to not-at-all. In some areas, science might provide a definitive answer to the question of whether an individual case is an instance of a general phenomenon. If tobacco smoke is the only cause of lung cancer, we logically know that someone with lung cancer got sick from tobacco smoke. In other areas, science might help increase the accuracy of individual decision-making along a range of helpfulness, from nearly determinative to just above random chance. If tobacco smoke causes lung cancer, but many other things, known and unknown, do so as well, we cannot say with certainty that the person’s lung cancer was caused by tobacco smoke. The degree of certainty that the science provides, of course, is the operative question. Indeed, sometimes even very good science will not demonstrably improve the accuracy of individual decision-making, though it might nonetheless be relevant and admissible because it provides the triers of fact with contextual information that will help them understand other evidence in the case.

A. When General Science Is Determinative in Particular Cases

In practice, the law is interested not simply in whether a particular variable causes a particular effect, but, ultimately, in whether a particular variable did cause the effect. Scientific research will sometimes identify a single unidirectional relationship between two variables. In medicine, the term pathognomonic refers to a diagnostic version of this insight. A

---

10 This analysis simplifies matters considerably, since both the existence and extent of the cause, as well as the existence and extent of the effect, may be disputed in a particular case.
symptom is pathognomonic when it is “decisively characteristic of a disease.””\textsuperscript{11} For example, “Koplik’s spots . . . are pathognomonic of measles.”\textsuperscript{12} The strongest version of a path-specific relationship would be the unusual situation where a cause and an effect are uniquely associated, such that the cause always produces the effect and the effect is always attributable to the cause. Outside of basic physics and chemistry, however, the strongest version of path-specificity will be quite rare. Nonetheless, such relationships are possible. This strong version could be termed cause/effect path-specificity because the cause and the effect are uniquely tied to one another.

The law is also interested in weaker versions of path-specificity. For instance, a particular cause might always produce a particular effect, but other causes might produce similar effects. This could be termed causal path-specificity because the cause always produces a single effect, but other causes might produce the same effect. An example of this might be a lesion in a specific part of the brain that produces auditory hallucinations. Anyone with such a lesion would suffer from auditory hallucinations, but not all people with auditory hallucinations have a lesion in that region of the brain. Conversely, a particular effect might always be produced by a particular cause, but the cause does not invariably produce the effect. This could be termed effect path-specificity because the effect has a single cause, but the cause does not have a single effect. An example of this is the relationship between asbestos exposure and mesothelioma. The unique cause of mesothelioma is exposure to asbestos, but not everyone exposed to asbestos develops mesothelioma.\textsuperscript{13}

In legal proceedings, the strength and nature of path-specificity is likely to be important. In general, cause/effect path-specificity will be the most probative kind of scientific evidence available. In contrast, the probative power of causal path-specificity or effect path-specificity will depend on the substantive law of the case. For example, in many criminal


\textsuperscript{13} Asbestos also causes other ailments, including lung cancer. See Piero Mustacchi, \textit{Lung Cancer Latency and Asbestos Liability}, 17 J. LEGAL MED. 277, 280 (1996). But, as mentioned, some people who are exposed to asbestos never get sick from it.
cases, the issue will be whether the defendant suffered the relevant effect, and it will not matter greatly that a variety of causes can produce it. In such cases, scientific evidence of causal path-specificity would strongly support the defendant's case. This would be so in an insanity case in which evidence that the defendant has a brain lesion that invariably produces auditory hallucinations would be highly probative, despite the fact that other factors might cause the same symptoms. Conversely, in many civil cases, effect path-specificity will be the more probative kind of evidence. In the example of mesothelioma, a civil plaintiff who has this disease will be able to trace it back to asbestos exposure. In many civil cases, a substantial obstacle to a plaintiff's recovery is showing that the effect he or she suffers from is attributable to the cause associated with the defendant. Effect path-specificity solves this difficulty. If the defendant was responsible for the plaintiff's asbestos exposure, then the plaintiff's mesothelioma is attributable to the defendant.

B. When General Science Is Probative, but Not Determinative, in Particular Cases

In most applied science contexts, path-specificity is not possible, either because it does not exist in actuality or because scientists’ methods are unable to identify those cases in which it does exist. In most areas of interest to the law, scientific research provides knowledge about cause and effect relationships generally, but will be only more or less determinate on the question of whether a specific instance of an effect is attributable to a specific cause, or that a specific cause contributed to a particular effect. In this vast domain, applied scientific research comes in myriad forms and its value for deciding individual cases varies greatly. In some situations, the science will be nearly definitive regarding a specific cause and effect relation and in others it will do little more than increase the likelihood that a relevant relationship exists slightly above chance.

As is true with the concept of path-specificity discussed in the previous section, indeterminate scientific research might be relevant in legal proceedings in three separately identifiable ways, regarding (1) effect only, because the cause is known (or can be assumed), (2) cause only, because the effect is known (or can be assumed), or (3) both cause and effect. As will become clear in the discussion that follows, the intended purpose for
which the science is to be used is associated with the demands
that courts place on the science itself.

In many legal contexts, only the effect is relevant
because the causal variable is fairly known or is assumed.
Indeed, one of the best known subjects in law and psychology
fits this category: eyewitness identification. In eyewitness
identification research, researchers have found that certain
factors interfere with accuracy, such as presence of a weapon,
cross-race identifications, and use of leading questions by
interviewers. In this example, the causal side of the equation
is the independent variable, which is more or less known or
assumed to be present in the case. The focus, therefore, is
principally on what effect this causal variable has had. Hence,
if the witness is white and the perpetrator is black, the
empirical crux of the matter concerns what effect this causal
variable has on the accuracy of the identification. Other
examples in which the effects are relevant and the cause is
known or assumed include the effects of hypnosis on memory,
the impact of putatively prejudicial photographs or images on
fact finders’ judgments, and the effect of violent television on
viewers.

In effect-relevant cases—that is, where the cause is
known or assumed and the effects have been the subject of
research—the science is rarely employed to do more than
provide general insights about those who have experienced the
causal variable of interest. It may very well be, for instance,
that when a gun is present, eyewitness identifications are on
average less accurate than when one is not; but this finding
provides very little information regarding whether any
particular identification is accurate. In the law, general
research findings might very well be relevant and admissible to
inform the jury of factors that might interfere with a witness’s
accurate recall, which the jury could use or ignore as it deemed

14 See Gary L. Wells, Eyewitness Identification: Scientific Status, in 2
Modern Scientific Evidence: The Law and Science of Expert Testimony, supra
note 7, at 520, 534-47 (Faigman et al. eds., 2009).
15 See Michael Nash & Robert Nadon, Hypnosis: Scientific Status, in 2
Modern Scientific Evidence: The Law and Science of Expert Testimony, supra
note 7, at 733.
16 David A. Bright & Jane Goodman-Delahunty, Gruesome Evidence and
17 Kevin D. Browne & Catherine Hamilton-Giachritsis, The Influence of
Violent Media on Children and Adolescents: A Public-Health Approach, 365 Lancet
702, 702 (2005).
fit. The science in this case, however, says very little about eyewitness identification.

The second category, and one that arises often in court, is when the effect is fairly known (or can be assumed), and the science is offered to demonstrate the cause of that effect. Whole areas of medical and psychological causation fit this category, as do some areas of forensic science. In medical causation, a plaintiff might be known to have leukemia (i.e., the effect) and the scientifically controverted issue will be whether one variable (e.g., trichloroethylene) or another caused the illness. In psychological causation, the same analysis applies. For example, a witness who suffers from Post-traumatic Stress Disorder (PTSD) might claim that it was caused by a sexual assault rather than other causes, such as a failed marriage and a lost job. Finally, some areas of forensic science fit this cause-relevant category. The best example is arson investigation. In the ordinary arson case, the effect is known (i.e., a burned or exploded structure), but the science is offered to demonstrate the cause (e.g., purposely set using some incendiary device or material).

When the proffered science is relevant to the cause of some known effect, it is ultimately meant to operate diagnostically in regard to the individual case at hand. This category presents the most classic manifestation of the challenges associated with reasoning from group data to decisions in individual cases. In many areas, the research provides substantial evidence of a general connection between variables, but the science does not pave a direct path to extrapolating from the general data to the individual case. Ordinarily, some additional method is used to bring the general science to the individual case, usually labeled vaguely as “differential diagnosis” or “differential etiology.” This issue is considered in Part II, infra.

The third and final category of scientific relevance is something of a catch-all, and involves those cases in which the science informs both the causal and the effect sides of the equation. In other words, in this category the situation or context is argued to have legal significance, but the science is necessary to show how or why this is so. Many psychological claims fall into this category, as do most of the forensic identification technologies. A good example of the former is

18 See generally Faigman et al., supra note 7, at 27–49.
research on predictions of violence. The matter of predicting violence has wide significance in the law, and scientists have sought to provide guidance on this issue by relating one set of variables (i.e., predictors) to another variable (i.e., future violence). Neither the “cause” nor the “effect” is known outside of the applicable research. Most forensic identification technologies operate similarly. Scientific research on DNA profiling, for instance, describes both the existence of the phenomenon as well as the significance of that phenomenon for legal decision-making. Significantly, both actuarial predictions of violence and DNA profiling are framed generally, and, to the extent that they are applied to individual cases, the proffered opinions ordinarily remain in their general population-based form.

In the end, law and science are separate disciplines and, though they often share goals or objectives, neither is nor should be expected to be the other’s handmaiden. It is hardly surprising, therefore, that the methods of science do not correspond neatly to the needs of the law. Yet, nonetheless, at least in a preliminary way, it is possible to identify general pathways of scientific investigation and consider how they sometimes might, but oftentimes do not, provide the answers to the questions the law poses. Understanding the parameters of the scientific enterprise, however, is only the first step in improving the law’s use of research data. Much of the information the law needs from science does not fit neatly into conventional modes of empirical inquiry. Whereas scientists ordinarily study causes and effects in populations, courts ordinarily need to determine causes and effects in particular individuals. The next section examines the difficulties endemic to developing a rigorous individual-based empiricism.

---

19 See Helena Kraemer et al., Coming to Terms With the Terms of Risk, 54 ARCHIVES GEN. PSYCHIATRY 337, 340 (1997). It should be noted that very often predictions of violence opinions are not based on scientific research at all. Many, if not most opinions offered in court on this subject, are based on clinical judgment, and they are presented accordingly as conclusions about a particular person. On the value of clinical versus actuarial predictions of violence, see Stefania Aegisdottir, The Meta-Analysis of Clinical Judgment Project: Fifty-Six Years of Accumulated Research on Clinical Versus Actuarial Predictions of Violence, 34 COUNSELING PSYCHOLOGIST 400 (2006).
II. FRAMING EMPIRICAL QUESTIONS IN THE COURTROOM

The basic perspective of most courtroom proceedings is individual and specific. Courts look to answer such questions as whether the defendant killed the victim, the plaintiff’s leukemia was caused by a chemical produced by the defendant, the juvenile defendant is competent to be tried as an adult, the capital defendant is likely to be violent if not executed, and so forth. While the ultimate issue in most legal proceedings involves the determination of a particular fact (or facts), courts well understand that underlying these specific questions is knowledge about the general world. Hence, a defendant’s guilt might depend on the general match probabilities of DNA evidence, and a plaintiff’s civil claim against a chemical manufacturer might depend partly on epidemiological studies showing an association between the alleged offending chemical and leukemia. Tackling the complex challenge of integrating scientific research into legal decision-making would be helped considerably if there were a vocabulary that permitted categorization of the different ways science might be relevant to legal decision-making. There has been no shortage of attempts at providing such a taxonomy.  

A. Taxonomies of Fact-Finding

The first, and still most influential, taxonomy of fact-finding in law was offered by Professor Kenneth Culp Davis. Davis distinguished between what he termed legislative facts and adjudicative facts. Legislative facts are those facts that transcend the particular dispute and are relevant to legal reasoning and the fashioning of legal rules. Adjudicative facts, in contrast, are those facts particular to the dispute. In a series of influential articles in the 1980s, Professors John Monahan and Laurens Walker refined Davis’ dichotomy in a manner that more fully captures the ways that science is

20 I too have participated in this endeavor, though my efforts were restricted to constitutional cases. See DAVID L. FAIGMAN, CONSTITUTIONAL FICTIONS: A UNIFIED THEORY OF CONSTITUTIONAL FACTS 43-62 (2008).
22 Id. at 402; see also Fed. R. Evid. 201(a), Advisory Committee’s Note (“Legislative facts . . . are those which have relevance to legal reasoning and the lawmaking process, whether in the formulation of a legal principle or ruling by a judge or court or in the enactment of a legislative body.”).
23 Davis, supra note 21, at 402.
used in the courtroom. Their primary focus was on the law’s use of social science. They identified three levels of convergence between social science and law: social authority, social facts, and social frameworks. Social authority refers to social science research relevant to the determination of legislative facts and thus the formulation of legal rules. According to their proposal, social authority is analogous to legal authority and should be consulted similarly. Hence, judges would consider social science “precedent” (i.e., past research) as presented through briefs, arguments, and sua sponte. The information found to be relevant and valid would then be incorporated into the judge’s conclusions of law. Alternatively, in the Monahan-Walker model, social science research might be relevant to adjudicative facts (what they call “social facts”), in which case, after being deemed admissible, it would be presented to the trier of fact through expert testimony. Finally, social science research might have relevance as a combination of social authority and adjudicative fact. Professors Monahan and Walker label this use “social frameworks,” where some issue in the particular dispute is claimed to be an instance of a social scientific finding or theory of general import.

The Monahan-Walker model, though framed to deal with their subject of interest (social science), nicely captures the three basic divisions of fact-finding that courts must process. Most importantly, their social framework category is a significant leap forward in clarifying the challenges associated with integrating empirical research into legal decision-making. Indeed, arguably the social authority (i.e., legislative facts) and social facts (i.e., adjudicative facts) are merely components of

---

28 Walker & Monahan (1987), supra note 24, at 563-67. According to Monahan and Walker’s social framework model, the judge would consider and instruct the jury on the accuracy of the general claim, but the jury would also hear expert testimony on how the research applies to the case before it. Id. at 592. In traditional practice, however, the jury is the fact finder for both components of social framework evidence. For present purposes, I need not choose which procedural approach is the better one.
social frameworks, with the latter two being defined as a function of the legal use for evidence, not its scientific nature. In other words, all empirical research is conceivable in terms of frameworks, because it invariably has both a general component and a specific component. Whether the general component is legally relevant at all and, if so, what it is relevant to prove, dictates in the Monahan-Walker model whether it is a “social authority” or “social framework.” For example, consider the empirical question of the developmental competence of sixteen- and seventeen-year-olds. In the context of capital punishment, this general fact was used in *Roper v. Simmons* to support the conclusion that applying the death penalty to those who killed before reaching the age of majority was unconstitutional. As such, this legislative fact was informed by “social authority.” On the other hand, if the question was whether a particular sixteen- or seventeen-year-old had competently waived his *Miranda* rights, the research used in *Roper* would be employed to inform a “social framework.” In the case involving the waiving of *Miranda* rights, the court would have to apply the framework to the individual case, thus paradigmatically using both components of Monahan and Walker’s social framework category.

**B. Empirical Frameworks**

For the purpose of examining evidentiary incommensurability between law and science, the Monahan and Walker concept of social frameworks is all that is specifically needed. It fully captures the juxtaposition of the inordinate empirical difficulties surrounding the use of group data to make individual decisions, and the law’s frequent need to do just that. Since the phenomenon of interest extends well beyond social science, and includes all applied science with policy implications, the term “empirical framework” is more accurate and will be used here. The following sections, therefore, consider the legal demands on empirical research, from both the more conciliatory use of general research data to answer general legal propositions, to the more demanding use of general data to reach individualized judgments.

---

C. Defining the “Frame”

Because ordinary science operates at the general level of descriptive and inferential statistics, it can be readily employed to determine general propositions. Consider, for example, a hypothesis that has been the subject of several legal cases: violent video games cause minors who play them to be violent and asocial. This hypothesis has been studied in a multitude of ways, including observational case studies, correlational studies, laboratory experiments, brain imaging, and so forth.31 If these differing methods point in the same direction, then some general conclusions might be made regarding the relationship between violent video games and violence among children. If they point in different directions, of course, the task is complicated greatly, if not made impossible, until more research is done. But even when the body of research is robust, conclusions are likely to be tentative and, at best, described in probabilistic terms.

The legal relevance of the science, however uncertainly known, depends on the substantive law of the case. In regard to the violent video game example, then, this hypothesis might be relevant as a general proposition—e.g., do violent video games lead to increased violence among children—or as that research might apply in a particular case—e.g., was the minor-defendant’s violent action attributable to having played violent video games.

In the law, most litigation tends to involve the application of general principles to a specific case. Frequently, however, a general proposition of science is itself at issue. A good example of this, coming from the violence in media example, is the case Entertainment Software Association v. Blagojevich.32 In Entertainment Software, several video industry trade associations sued the State to enjoin the enforcement of two statutes that regulated the content of violent and sexually explicit videos. The plaintiffs argued that the State’s laws violated the Free Speech Clause of the First Amendment. The district court agreed that the laws implicated First Amendment rights and held that the legislation could survive only if the State had a compelling interest that would be substantially achieved by the laws. The court found that

“[t]he Illinois General Assembly’s main justifications . . . were three legislative findings about the effect of playing video games on minors’ physiological and neurological development.” According to the court, the legislature believed that playing violent video games makes children (1) “exhibit violent, asocial, or aggressive behavior”; (2) “[e]xperience feelings of aggression”; and (3) “[e]xperience a reduction of activity in the frontal lobes of the brain which is responsible for controlling behavior.” In concluding that Illinois had not met its considerable burden, the court extensively reviewed psychological and neurological research that had been advanced by the State. The court explained that the State “failed to present substantial evidence showing that playing violent video games causes minors to have aggressive feelings or engage in aggressive behavior.” Moreover, the court stated that “there is barely any evidence at all, let alone substantial evidence, showing that playing violent video games causes minors to experience a reduction of activity in the frontal lobes of the brain which is responsible for controlling behavior.” The court permanently enjoined the Illinois law.

The second hypothesis, that a particular minor’s violent action is attributable to having played violent video games, is the more typical courtroom situation in regard to scientific evidence. In these cases, both the general hypothesis and the specific hypothesis are at issue. Although the defense is unusual, defendants have on occasion argued insanity on the basis of video programming. In Zamora v. State, for example, “Zamora’s insanity defense was based upon ‘involuntary subliminal television intoxication.’” In particular, defense counsel argued that violent television had a noxious influence on sociopathic children and that Zamora had killed as a consequence of this effect. To support this theory, the defense

---

33 Id. at 1073.
34 Id.
35 Id. at 1074 (The court added that, “[a]t most, researchers have been able to show a correlation between playing violent video games and a slightly increased level of aggressive thoughts and behavior.”).
36 Id.
39 Id.
offered two experts. The first, a psychologist, offered to testify to the effect of television on adolescents generally. A second expert, a psychiatrist, testified that the defendant “did not know right from wrong” when he “fired the fatal shot,” thus applying the general theory of the case to the particular defendant. The court excluded the psychologist on the ground that she could not speak to Zamora’s individual case. The psychiatrist testified at trial, but apparently to little effect, since Zamora was convicted.

In the courtroom, research on general propositions, such as whether violent media causes an increase in violence among children, addresses a threshold question and one which scientists are trained to address. In an insanity defense to murder, however, the question is whether the particular person’s violence was caused by exposure to violent media. This issue of specific application poses a complex and difficult cognitive exercise. Moreover, it is an exercise that varies in different empirical contexts. It is also a subject that has been substantially ignored by scientists interested in the courtroom use of their data.

D. Reasoning to the Specific

Although the challenge of reasoning from general research data to individual cases has been considered in a fairly cursory manner by courts and legal scholars, the basic challenges are fairly easily summarized. This is especially so in the conventional toxic tort litigation context, the area in which courts have most often considered it. In a nutshell, the first task is to demonstrate that the substance could have caused the ailment (i.e., the validity of the general proposition); the second task is to show both that it probably did, and that other substances probably did not, cause the plaintiff’s condition.

The simplest case of this reasoning process might involve general research that indicates that some substance causes an ailment that is uniquely associated with that substance. For instance, as noted in Part I, asbestos has been shown to cause mesothelioma and it is the only substance known to cause it. Since mesothelioma is a “signature disease,” the only question concerns the circumstances of the

40 Id.
individual's exposure to asbestos (i.e., was the defendant responsible), not whether exposure caused the condition. The cause and effect path-specificity operates in this example to permit straightforward logical deductions from the general data to individual cases. This is rare in toxic tort litigation. For example, in contrast to asbestos, while second-hand smoke has been linked to lung cancer, many other substances are known to cause lung cancer. Hence, in regard to identifying the cause of a person's lung cancer, an expert must not only rule-in smoking as a possible cause, but also rule-out other possible causes. 42

The principal tool used to move from general research findings to statements about individual cases is “differential etiology,” sometimes misleadingly referred to as “differential diagnosis.” Properly understood, differential diagnosis refers to the identification of the illness or behavioral condition that a person is experiencing. Differential etiology refers to the cause or causes of that condition. Hence, the determination that a person suffers from “dissociative amnesia” and not “dissociative fugue” is a diagnostic issue. 43 The determination that a sexual assault at age ten caused the diagnosed dissociative amnesia, and that it did not result from a medical condition or physical trauma, is an etiological matter. Very different skill sets are usually involved in these two determinations. Indeed, the DSM explicitly eschews any claim of the etiological verity of its diagnostic categories. 44 It is worth emphasizing, as well, that the validity of the diagnosis of dissociative amnesia is a matter of general research. The entire process of differential diagnosis and differential etiology assumes that the designated category has adequate empirical support in the first place as a general proposition. Hence, although it is logically obvious, it should be stated plainly that an expert should never be permitted to testify about a specific application of a general proposition if research does not adequately support the general proposition.

In the professional practice of both clinical medicine and clinical psychology, the primary concern is diagnosis and not etiology. An oncologist might be curious about what caused his or her patient’s leukemia, but the doctor’s first task is to

44 Id. at xxxvii.
diagnose and treat the condition, not determine whether it was caused by trichloroethylene, benzene, electromagnetic fields, or something else. Similarly, a psychologist treating a person thought to suffer from either Post-traumatic Stress Disorder (PTSD) or adjustment disorder is primarily concerned with identifying and treating the condition, not determining the true causes of that condition. In the ordinary practice of clinical medicine and clinical psychology, treatment and therapy are the principal objectives, not assessing cause. A person presenting symptoms associated with PTSD, therefore, may claim that the traumatic event was a sexual assault committed by her uncle. From the therapeutic standpoint, at least at the start, the important factor is that there was a traumatic event. Whether the patient's uncle was the cause need not be specifically resolved for diagnostic purposes. In the law, of course, who caused the traumatic event is the crux of the matter. Hence, the core nature of much clinical practice is at right angles to the crux of most legal inquiries.

In the courtroom, differential etiology is the operative issue. Moreover, the same basic principle is implicated, whether the expert opinion comes from research-based science or clinical practice (i.e., "experience"). Indeed, at least superficially, the former suffers a comparative disadvantage, since the research tradition does not ordinarily purport to offer conclusive statements about individual cases. Research, for example, might identify factors highly associated with false confessions, but these general propositions are some distance from what is needed to allow experts to opine regarding the truth or falsity of any particular confession. Clinicians at least have a history of applying general knowledge to individual cases, though, as noted, while this practice might be well accepted for therapeutic purposes, its validity for forensic ends is somewhat doubtful. Whether researchers or clinicians have the wherewithal to help triers of fact in applying general research propositions to specific cases is a threshold legal matter that should depend on the reliability and validity of the differential etiology done in the respective case. It may be, that is, that in vast areas of clinical practice there is no general research foundation in the first instance. And, as stated above, if research does not support a general proposition—say, the phenomenon of repressed memories—then clinical expert testimony that a particular person has repressed certain memories of early sexual abuse cannot be sustained.
E. Differential Etiology

Differential etiology is a reasoning process that involves a multitude of factors, few of which are easily quantified. An expert offering an opinion regarding a specific case must first consider the strength of the evidence for the general proposition being applied in the case. If the claim is that substance X caused plaintiff's condition Y, the initial inquiry must concern the strength of the relationship between X and Y as a general proposition. For example, both second-hand smoke and first-hand smoke are associated with lung cancer, but the strength of the relationship generally is much stronger for the latter than it is for the former. The inquiry regarding strength of relationship will depend on many factors, including, among other things, the statistical strength of any claims and the quality of the methods used in the research. Additionally, the general model must consider the strength of the evidence for alternative possible causes of Y and the strength of their respective relationships (and possibly interactions with other factors). Again, the quality of the research and the different methodologies employed will make comparisons difficult. Complicating matters further regarding identification of potential causes of condition Y are the myriad of possible causes that have not been studied, or have been studied inadequately.\(^45\) Hence, determining the contours of the general model is a dicey affair in itself, since it requires combining disparate research results and discounting those results by an unknown factor associated with additional variables not yet studied. And this is just the first part of the necessary analysis if the expert wants to give an opinion about an individual case.

The second part of the analysis—specific application of general propositions that are themselves supported by adequate research—requires two abilities, neither of which are clearly within most scientists’ skill sets. The first, and perhaps less problematic, is that of forensic investigator. Almost no matter what the empirical relationship, whether medical or psychological, exposure or dosage levels will be relevant to the diagnosis. The first principle of toxicology is that “the dose

\(^{45}\) In Henricksen v. Conocophils Co., 605 F. Supp. 2d 1142 (E.D. Wash. 2009), the court observed that eighty to ninety percent of the causes of acute myelogenous leukemia (AML) were unknown (“ideopathic”). Id. at 1149. The court stated that “[i]f 90 percent of the causes of a disease are unknown, it is impossible to eliminate an unknown disease as the efficient cause of a patient’s illness.” Id. at 1162 (quoting Whiting v. Boston Edison Co., 891 F. Supp. 12, 21 n.41 (D. Mass. 1995)).
makes the poison,” since any substance in sufficient quantities
could injure or kill someone." Similarly, in a wide variety of
psychological contexts, the exposure or dose will be the poison.
For instance, degree of trauma affects diagnostic categorization
between PTSD and adjustment disorder, level of anxiety affects
eyewitness identifications, amount of lack of sleep affects false
confession rates, and so on. The expert testifying to specific
causation must determine exposure and dosage levels for the
suspected cause (i.e., the source suspected by the client) as well
as for all other known or possible causes. This task is difficult
enough alone, but is enormously complicated by the significant
potential for recall bias, given that the litigation will be
profoundly affected by what is recalled.

The second skill set that is needed has not yet been
invented or even described with precision. Somehow, the
diagnostician must combine the surfeit of information
concerning the multitude of factors that make up the general
model, combine it with the case history information known or
suspected about the individual, and offer an opinion with some
level of confidence that substance or experience X was the
likely cause of condition Y. In practice, this opinion is usually
stated as follows: “Within a reasonable degree of
medical/psychological certainty, it is my opinion that X caused
[a particular case of] Y.” This expression has no empirical
meaning and is simply a mantra repeated by experts for
purposes of legal decision makers who similarly have no idea
what it means. But even less extreme versions of this
statement—such as, “It is more likely than not true that this
case is an instance of some general phenomenon”—are
objectionable. Just how, for instance, would an eyewitness
researcher determine that a witness was more likely than not
inaccurate when the witness made a cross-racial identification
of the defendant after seeing the unarmed perpetrator for five
minutes under a streetlight from an unobstructed view twenty
feet away from the crime? There are no data that would
support psychologists’ ability to make such statements,
however modest or innocuous they may appear. Experts’ case-
specific conclusions appear to be based largely on an admixture
of an unknown combination of knowledge of the subject,

---

46 Bernhard D. Goldstein & Russellyn Carruth, Toxicology: Scientific Status,
in 3 Modern Scientific Evidence: The Law and Science of Expert Testimony, supra
note 7, at § 22.
experience over the years, commitment to the client or cause, intuition, and blind-faith. Science it is not.

Whether, and in what way, particular scientific findings are relevant to legal decision-making depends on the substantive law of the case. Frequently, the relevant factual issue under applicable law involves general propositions, ones that population-based research corresponds to directly. Much more often, however, the empirical focus of the ultimate legal issue is on the particular case. But conventional scientific methods do not share this focus. Although research data might demonstrate with high confidence that a particular variable has an effect of interest, it typically cannot demonstrate with the same confidence that the particular variable had the effect of interest in a particular case. Reconciling this evidentiary incommensurability between what science ordinarily does and what the law ordinarily needs is, as yet, one of the great unmet challenges at the intersection of science and the law.

CONCLUSION

Most evidentiary codes require that expert testimony “assist the trier of fact” in order for it to be admissible.\(^\text{47}\) Scientific expert testimony, however, must be legally relevant and have evidentiary reliability (i.e., scientific validity).\(^\text{48}\) Moreover, expert opinion must offer insights beyond what triers of fact could do on their own. Put another way, scientist-experts are limited to testifying about what their respective field’s research can validly add to fact-finders’ deliberations—and nothing more. This injunction, however, is not always followed. In particular, experts frequently seek to comment not simply on the import of general research findings, but on whether a particular case fits those findings. Scientific research that permits a valid description of a general phenomenon, however, does not invariably give experts the capacity to validly determine whether an individual case is an instance of that general phenomenon.

A basic difference in perspective between science and the law is that science studies individuals in order to make statements about populations, while the law studies populations in order to make statements about individuals. It

---

\(^\text{47}\) Fed. R. of Evid. 702.

does not necessarily follow that a scientist who can validly describe a general phenomenon also has the wherewithal to say whether an individual case is an instance of that general phenomenon. In many respects, the matter of translating scientific research findings into helpful information for fact-finders in court should be a subject of first concern for applied science. Yet this issue has been largely ignored by scientists. This essay calls for a broadly conceived collaborative effort to consider this basic issue, one that is endemic to the intersection of law and science.
Scientific Evidence in Criminal Prosecutions

A RETROSPECTIVE

Paul C. Giannelli

The publication of the National Academy of Sciences (NAS) Report on forensic science, Strengthening Forensic Science in the United States: A Path Forward,1 in February 2009 marked the culmination of thirty years of debate on the admissibility of scientific evidence. In a sense, the NAS Report told Congress to scrap the current structure and replace it with a system that was independent of law enforcement and premised on the research norms of science.2 The impetus for the report can be traced to two events: The Supreme Court’s decision in Daubert v. Merrell Dow Pharmaceuticals, Inc.,3 an opinion that revolutionized the legal test for the admissibility of expert testimony, and DNA analysis, a technique that

---

1 Albert J. Weatherhead III & Richard W. Weatherhead Professor of Law, Case Western Reserve University. Like most evidence teachers, I am deeply indebted to Margaret Berger. I began teaching evidence in 1975, the year the Federal Rules of Evidence became effective. In preparing for class, I relied on Jack B. Weinstein & Margaret A. Berger, Weinstein’s Evidence, which was the only complete text on the Federal Rules at the time. In class, I used Jack B. Weinstein, John Mansfield, Margaret A. Berger & Norman Abrams, Cases and Materials on Evidence, as my casebook. See Paul C. Giannelli, Book Review: Cases and Materials on Evidence, 49 Brook. L. Rev. 629, 633-34 (1983) (“In summary, the seventh edition improves what was already an exceptional book. The comprehensiveness of the text, achieved in part through the use of copious notes, has been retained, and the organizational changes will assist in the effective presentation of the course material.”).


revolutionized forensic science. Professor Margaret Berger played a significant role in both these developments, as well as in the NAS Report itself.

I. THE DAUBERT TRILOGY

The Federal Rules of Evidence became effective in 1975. At that time, the leading case on the admissibility of scientific evidence was Frye v. United States, which held that the admissibility of expert testimony depended on its “general acceptance in the particular field in which it belongs.” In 1974, the D.C. Circuit observed that Frye had “been followed uniformly in this and other Circuits and there has never been any successful challenge to it in any federal court.” Frye was also the majority rule in the states. Yet, neither Frye nor the admissibility of novel scientific evidence was addressed in the legislative history of the Federal Rules. The issue was ignored in the advisory committee’s notes, the congressional committee reports, and the extensive hearings on the Federal Rules. The year before the Supreme Court decided Daubert, Judge Becker and Professor Orenstein referred to the Frye issue as the “most controversial and important unresolved question in the Federal Rules” of Evidence.

---

5 Frye, 293 F. at 1014.
6 United States v. Skeens, 494 F.2d 1050, 1053 (D.C. Cir. 1974); see also United States v. Alexander, 526 F.2d 161, 163 n.3 (8th Cir. 1975) (The “federal courts of appeals continue to subscribe to [the] ‘general scientific acceptability’ criterion.”).
7 See Reed v. State, 391 A.2d 364, 368 (Md. 1978) (“This criterion of ‘general acceptance’ in the scientific community has come to be the standard in almost all of the courts in the country which have considered the question of the admissibility of scientific evidence.”).
During this pre-Daubert era, textwriters\textsuperscript{12} and law review commentators\textsuperscript{13} disagreed sharply about the continued viability of the Frye test under the Federal Rules. Moreover, if Frye was discarded, what would replace it? Arguing that the Federal Rules’ failure to incorporate Frye indicated its abandonment, Judge Weinstein and Professor Berger proposed an alternative approach specifying a number of factors for determining the reliability of expert testimony: (1) the new technique’s general acceptance in the field, (2) the expert’s qualifications and stature, (3) the use that has been made of the technique, (4) the potential rate of error, (5) the existence of specialized literature, (6) the novelty of the new invention, and (7) the extent to which the technique relies on the subjective interpretation of the expert.\textsuperscript{14} At a conference devoted to this issue, Professor Berger argued for this enhanced reliability test,\textsuperscript{15} commenting that the “Frye test often seems to obscure what the lawyers really should be asking. The question is not always whether a procedure is scientifically valid, but whether the procedure is being applied appropriately under the circumstances in a particular case.”\textsuperscript{16} At another point, she reported for a breakout group:


\textsuperscript{14} 3 Weinstein et al., supra note 12, at ¶ 702[03].


\textsuperscript{16} 99 F.R.D. at 222.
No one in our group thought that the general acceptance test should be retained, because no one could state specifically just what it means. On the other hand, we do not think that all scientific evidence should be admitted under a loosely structured relevance standard. We agreed that the court should conduct some sort of preliminary screening to ensure that a threshold of validity has been met.\(^1\)

The Supreme Court’s approach in *Daubert* echoed Professor Berger’s position in several respects. First, the Court jettisoned *Frye* as a matter of statutory interpretation. Second, the Court required an independent judicial assessment of the reliability of expert testimony. Third, the Court’s reliability test rested on a multi-factor analysis, albeit with some factors that differed from the Weinstein-Berger proposal.\(^2\) In the aftermath of *Daubert*, many evidence scholars attempted to predict the ramifications of the decision.\(^2\) Few were as perceptive as Professor Berger, who wrote on the topic the year after *Daubert* was decided.\(^2\) She made three observations about expert testimony in criminal cases.

### A. Lack of Empirical Research

Her first point stressed the lack of empirical research: “Considerable forensic evidence made its way into the courtroom without empirical validation of the underlying theory and/or its particular application. Courts never required some of the most venerable branches of forensic science—such as fingerprinting, ballistics, and handwriting—to demonstrate their ability to make unique identifications.”\(^2\) A year later, challenges to the admissibility of these forensic techniques began. The first significant challenge under *Daubert* involved handwriting and came in *United States v. Starzecepyzel*,\(^2\)

---

\(^1\) *Id.* at 230.

\(^2\) In describing the trial judge’s screening or “gatekeeping function,” the Court identified a number of factors: (1) testability, (2) peer review and publication, (3) error rate, (4) maintenance of standards, and (5) general acceptance. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 593 (1993).

\(^3\) *See generally Symposium, Scientific Evidence After the Death of Frye*, 15 *CARDOZO L. REV.* 1745 (1994).


\(^5\) *Id.* at 1354.

decided in 1995. Other handwriting cases followed, pointing out the lack of empirical support underpinning the technique. Significantly, these cases viewed the Daubert trilogy as inviting “reexamination even of ‘generally accepted’ venerable, technical fields.”

If Starzecpyzel unnerved document examiners, United States v. Llera Plaza “sent shock waves through the community of fingerprint analysts.” In that case, Judge Pollak ruled that fingerprint experts would not be permitted to testify that two sets of prints “matched”—that is, a positive identification to the exclusion of all other persons. This was the first time in nearly a hundred years that such a decision had been rendered. On rehearing, however, Judge Pollak reversed

\[23\] Handwriting was a prime target because a comprehensive article questioning the underpinnings of the technique had been published in 1989. According to the authors of that article:

Our literature search for empirical evaluation of handwriting identification turned up one primitive and flawed validity study from nearly 50 years ago, one 1973 paper that raises the issue of consistency among examiners but presents only uncontrolled impressionistic and anecdotal information not qualifying as data in any rigorous sense, and a summary of one study in a 1978 government report. Beyond this, nothing.


\[24\] See, e.g., United States v. Hidalgo, 229 F. Supp. 2d 961, 967 (D. Ariz. 2002) (“Because the principle of uniqueness is without empirical support, we conclude that a document examiner will not be permitted to testify that the maker of a known document is the maker of the questioned document. Nor will a document examiner be able to testify as to identity in terms of probabilities.”); United States v. Lewis, 220 F. Supp. 2d 548, 554 (S.D. W. Va. 2002) (Expert’s “bald assertion that the ‘basic principle of handwriting identification has been proven time and time again through research in [his] field,’ without more specific substance, is inadequate to demonstrate testability and error rate.”).

\[25\] United States v. Hines, 55 F. Supp. 2d 62, 67 (D. Mass. 1999) (handwriting comparison); see also Hidalgo, 229 F. Supp. 2d at 966 (same) (“Courts are now confronting challenges to testimony, as here, whose admissibility had long been settled.”).


\[28\] The first reported fingerprint case was decided in 1911. See People v. Jennings, 96 N.E. 1077 (Ill. 1911). As Professor Mnookin has noted, however, “fingerprintes were accepted as an evidentiary tool without a great deal of scrutiny or skepticism.” Jennifer L. Mnookin, Fingerprint Evidence in an Age of DNA Profiling, 67 BROOK. L. REV. 13, 17 (2001). She elaborated: “Even if no two people had identical sets of fingerprints, this did not establish that no two people could have a single identical print, much less an identical part of a print. These are necessarily matters of
himself, and later cases would continue to uphold the admissibility of fingerprint evidence. Yet, the case had captured the attention of the media, with news reports, mainstream publications, scientific journals, and television shows giving it substantial coverage. Legal articles followed.

probability, but neither the court in Jennings nor subsequent judges ever required that fingerprint identification be placed on a secure statistical foundation." Id. at 19.


with many commentators believing that *Llera Plaza I* was more faithful to *Daubert* than *Llera Plaza II*.36

Once *Daubert* challenges on the admissibility of handwriting and fingerprint evidence had been filed, it was inevitable that firearms (ballistics) and tool mark identifications would also be questioned.37 Although the initial attacks failed,38 a pair of decisions by federal district courts in Boston changed all this. The first case, *United States v. Green*, was a frontal attack on the lack of empirical testing in this field. The expert testified that a match could be made “to the exclusion of every other firearm in the world.”39 That conclusion, according to Judge Gertner, was “extraordinary, particularly given [the expert’s] data and methods.”40 Consequently, the expert would only be permitted to explain the ways in which the casings were similar, but not that they came from a specific weapon “to the exclusion of every other

---


37 Firearms identification (“ballistics”) developed in the early part of the last century, and, by 1930, courts were admitting evidence based on the technique. Subsequent cases have followed these precedents, admitting evidence of bullet, cartridge case, and shot shell identifications. See 1 GIANNELLI & IMWINKELRIED, supra note 22, ch. 14 (discussing the scientific and legal issues associated with firearms and tool mark identifications).

38 See *United States v. Hicks*, 389 F.3d 514, 526 (5th Cir. 2004) (ruling that “the matching of spent shell casings to the weapon that fired them has been a recognized method of ballistics testing in this circuit for decades”); *United States v. Foster*, 300 F. Supp. 2d 375, 376 n.1 (D. Md. 2004) (“Ballistics evidence has been accepted in criminal cases for many years. . . . In the years since *Daubert*, numerous cases have confirmed the reliability of ballistics identification.”); *United States v. Santiago*, 199 F. Supp. 2d 101, 111 (S.D.N.Y. 2002) (“The Court has not found a single case in this Circuit that would suggest that the entire field of ballistics identification is unreliable.”); *State v. Anderson*, 624 S.E.2d 393, 398 (N.C. Ct. App. 2006) (no abuse of discretion in admitting bullet identification evidence); *Commonwealth v. Whitacre*, 878 A.2d 96, 101 (Pa. Super. Ct. 2005) (“no abuse of discretion in the trial court’s decision to permit admission of the evidence regarding comparison of the two shell casings with the shotgun owned by Appellant”).


40 Id. Although the expert had seven years of experience in the field, he was not certified, and his lab was not accredited. Moreover, he had never been formally tested by a neutral proficiency examination. Finally, he could not cite any reliable error rates. The expert “conceded, over and over again, that he relied mainly on his subjective judgment. There were no reference materials of any specificity, no national or even local database on which he relied. And although he relied on his past experience with these weapons, he had no notes or pictures memorializing his past.” Id.
firearm in the world.” In the court’s view, that conclusion “stretche[d] well beyond [the expert’s] data and methodology.” The court also issued a caution: “The more courts admit this type of toolmark evidence without requiring documentation, proficiency testing, or evidence of reliability, the more sloppy practices will endure; we should require more.”

The second case, United States v. Monteiro, resulted in a six-day evidentiary hearing. Although the court found that “the underlying scientific principle behind firearm identification—that firearms transfer unique toolmarks to spent cartridge cases—is valid under Daubert,” the expert in that case had yet to satisfy the other Daubert factors. More importantly, the court described the traditional methodology as essentially “tautological,” entrusting the critical decision to “the minds eye of the examiner.”

The next year, in United States v. Williams, the Second Circuit upheld the admissibility of firearms identification evidence, while noting that it did “not wish this opinion to be taken as saying that any proffered ballistic expert should be routinely admitted.” Moreover, the court observed that the

---

41 Id. at 109.
42 Id.
43 Id.
45 Id. at 355.
46 Because the expert did not make any sketches or take any photographs, adequate documentation was lacking: “Until the basis for the identification is described in such a way that the procedure performed by Sgt. Weddleton is reproducible and verifiable, it is inadmissible under Rule 702.” Id. at 374. Moreover, an independent second examiner had not confirmed the identification, which was particularly important because replacement parts had been used in the test-firing. Id.
47 The court wrote:

[The AFTE Theory, upon which the government relies, is tautological: it requires each examiner to decide when there is “sufficient agreement” of toolmarks to constitute an “identification.” . . . This threshold is surpassed when the examiner finds that the agreement of toolmarks “exceeds the best agreement demonstrated between toolmarks known to have been produced by different tools and is consistent with agreement demonstrated by toolmarks known to have been produced by the same tool.” . . . Toolmark analysis does not follow an objective standard requiring, say, a certain percentage of marks to match.

Id. at 370 (citations omitted) (citing Theory of Identification, Association of Firearm and Toolmark Examiners, 30 AFTE J. 86 (1998)).
48 506 F.3d 151 (2d Cir. 2007).
49 Id. at 161; see also United States v. Natson, 469 F. Supp. 2d 1253, 1261 (M.D. Ga. 2007) (“According to his testimony, these toolmarks were sufficiently similar to allow him to identify Defendant’s gun as the gun that fired the cartridge found at the crime scene. He opined that he held this opinion to a 100% degree of certainty. . . .
Daubert trilogy did not “‘grandfather’ or protect from Daubert scrutiny evidence that had previously been admitted under Frye.” Some trial courts continued to limit the scope of the testimony, finding that the record did not support the conclusion that identifications can be made to the exclusion of all other firearms in the world. Another court ruled that the expert would be permitted to testify only that it was “more likely than not” that recovered bullets and cartridge cases came from a particular weapon.

Other techniques, such as bite mark comparison, microscopic hair examination, bullet lead analysis, and intoxication testing, were also challenged. As Professor Berger had noted, the lack of empirical research was the critical issue.
B. Background Rates

Professor Berger’s second point focused on the most significant (and fundamental) problem underlying forensic identification expert testimony—i.e., the inherent probabilistic quality of these opinions. She wrote:

Prior to Daubert, courts admitted scientific evidence without noticing that, in some instances, the probative value of the evidence depends on background statistical information. If, for example, the samples of tape to which a defendant had access at his place of work match samples of tape used to manufacture a bomb sent through the mails from an unknown location, the probative value of that evidence is virtually non-existent if thousands of identical rolls of tape were distributed throughout the world. The crucial scientific inquiry in these cases is not only whether the technique is capable of producing matches, but also the probability that other matches exist.

. . . We allow eyewitnesses to testify that the person fleeing the scene wore a yellow jacket and permit proof that a defendant owned a yellow jacket without establishing the background rate of yellow jackets in the community. Jurors understand, however, that others than the accused own yellow jackets. When experts testify about samples matching in every respect, the jurors may be oblivious to the probability concerns if no background rate is offered, or may be unduly prejudiced or confused if the probability of a match is confused with the probability of guilt, or if a background rate is offered that does not have an adequate scientific foundation.

In 2008, a year before the NAS Report on forensic science was issued, a different NAS Report, one on computerized ballistic imaging, echoed Professor Berger’s point. This Report cautioned: “Conclusions drawn in firearms identification should not be made to imply the presence of a firm statistical basis when none has been demonstrated.” In particular, the authors of the Report were concerned about testimony cast “in bold absolutes” such as an assertion that a match can be made to the exclusion of all other firearms in the world. “Such comments cloak an inherently subjective assessment of a match with an extreme probability statement

---

reference to the same then-unchallenged authority. . . . I cannot agree that [various intoxication] tests, singly or in combination, have been shown to be as reliable as asserted by Dr. Burns, the NHTSA publications, and the publications of the communities of law enforcement officers and state prosecutors.”).

57 Procedural Paradigms, supra note 20, at 1356-57 (footnote omitted).
that has no firm grounding and unrealistically implies an error rate of zero.”

C. Error Rates

Professor Berger’s third point concerned one of the Daubert Court’s reliability factors, i.e., error rates. She noted that “[p]re-Daubert courts often ignored laboratory and technician error rates resulting from the subjectivity involved in interpreting particular forensic tests and overlooked the lack of proper laboratory procedures that can produce other kinds of errors.” As it turned out, in many post-Daubert cases, fingerprint experts testified that the “error rate for the method is zero.” Experts argued that, while individual examiners may make mistakes, the method itself is perfect. However, in this context the dichotomy between “methodological” and “human” error rates is “practically meaningless” because the examiner is the method and the examiner’s judgment is subjective.

In sum, a year after Daubert was decided, Professor Berger predicted the nature of the challenges that would be mounted in the forensic identification cases, and more importantly, identified the critical issues—opinion testimony masking probabilistic assumptions and subjective judgments without acknowledging error rates.

59 Id.
60 Procedural Paradigms, supra note 20, at 1358.
61 United States v. Havvard, 117 F. Supp. 2d 848, 854 (S.D. Ind. 2000), aff’d, 260 F.3d 597 (7th Cir. 2001). But see United States v. Mitchell, 365 F.3d 215, 246 (3d Cir. 2004) (“Testimony at the Daubert hearing indicated that some latent fingerprint examiners insist that there is no error rate associated with their activities or that the examination process is irreducibly subjective. This would be out-of-place under Rule 702.”).
62 See Mnookin, supra note 28, at 60. She goes on to provide this analogy: “The same argument could be made of eyewitness testimony, a notoriously unreliable form of evidence. People are all distinct from one another in observable ways; therefore the theoretical error rate of eyewitness identification is zero, though in practice observers may frequently make errors.” Id.
63 See Sandy L. Zabell, Fingerprint Evidence, 13 J.L. & Pol’y 143, 172 (2005) (“But, given its unavoidable subjective component, in latent print examination people are the process.”). In 2005, Professor Cole published an article documenting twenty-three cases of fingerprint misidentifications. See Simon A. Cole, More Than Zero: Accounting for Error in Latent Fingerprint Identification, 95 J. Crim. L. & Criminology 985, 999 (2005). The misidentification cases include some that involved (1) verification by one or more other examiners; (2) examiners certified by the International Association of Identification; (3) procedures using a sixteen-point standard; and (4) defense experts who corroborated misidentifications made by prosecution experts. Id. at passim.
II. DNA Evidence

The advent of DNA profiling in 1985 produced a sea change in forensic science. One court called DNA evidence the "single greatest advance in the search for truth . . . since the advent of cross-examination." Even its early critics acknowledged that "[a]ppropriately carried out and correctly interpreted, DNA typing is possibly the most powerful innovation in forensics since the development of fingerprinting in the last part of the 19th Century." No other technique had been as complex or so subject to rapid change. New DNA technologies were introduced at the trial level as cases litigating the older procedures worked their way through the appellate court system.

Although the introduction of DNA evidence went smoothly in the early going, a significant challenge to admissibility was mounted in People v. Castro, a 1989 case.


67 The initial technique, Restriction Fragment Length Polymorphism (RFLP) analysis by gel electrophoresis, was soon supplanted by Polymerase Chain Reaction (PCR)-based methods involving the DQ-alpha locus, "polymarkers," and the D1S80 locus. These, in turn, were replaced by Short Tandem Repeats, the current procedure. In addition to nuclear DNA analysis, courts have admitted evidence based on mitochondrial DNA (mtDNA) sequencing, as well as DNA analyses of animals, plants, and the HIV virus. See United States v. Boswell, 270 F.3d 1200 (8th Cir. 2001) (in a false statement prosecution, DNA used to compare swine blood); State v. Bogan, 905 P.2d 515 (Ariz. Ct. App. 1995) (in murder case, DNA of seed pods from palo verde trees at scene compared to those found in Bogan's truck); State v. Schmidt, 699 So. 2d 448 (La. Ct. App. 1997) (in a case of attempted murder by injection of HIV, expert testified that the strands of HIV from two persons were "closely related"). But see State v. Leuluiaii, 77 P.3d 1192 (Wash. Ct. App. 2003) (canine DNA match between sample obtained from defendant and murder victim's dog not generally accepted).

68 545 N.Y.S.2d 985 (N.Y. Sup. Ct. 1989). In an unusual occurrence, the prosecution and defense experts met without the attorneys and issued a joint statement, including the following: "[T]he DNA data in this case are not scientifically reliable enough to support the assertion that the samples . . . do or do not match. If these data were submitted to a peer-reviewed journal in support of a conclusion, they
The initial DNA skirmishes were over laboratory protocols, as in *Castro*, but the controversy quickly metamorphosed into fights over statistical interpretation and population genetics. Population geneticists used statistical techniques to define the extent to which a match of DNA markers individuated the accused as the source of the crime scene sample and were able to point to extensive empirical testing to support their opinions. The validity of the statistical methods became the focus of litigation.

As the dispute heated, the FBI requested the National Academy of Sciences to review the procedure. That organization issued two reports on the subject, noting the importance of certain practices: “No laboratory should let its results with a new DNA typing method be used in court, unless it has undergone . . . proficiency testing via blind trials.” The first NAS Report on DNA, however, provoked its own controversy and a second report was requested. The controversy centered on a proposal (the ceiling principle) offered to resolve the statistical issues surrounding DNA testimony. Professor Berger served on the second NAS Committee, and its report settled many of the controverted issues—and, as she later noted, “DNA profiling . . . is undoubtedly our ‘gold standard’ of expertise.”

Professor Berger subsequently wrote on other DNA issues. One article addressed the reporting of laboratory error
rates, an important but contentious issue. In another article, she examined the impact of DNA exonerations on the criminal justice system. In a book chapter, she considered DNA profiling’s impact on finality principles.

III. NAS FORENSIC SCIENCE REPORT

As noted at the beginning of this essay, the National Academy of Sciences issued its landmark report on forensic science in the beginning of 2009. Implementation of its recommendations would be the most important development in forensic science since the establishment of the crime laboratory in the mid-1920s. The issues are pressing. As the Report recognized, “[a]mong existing forensic methods, only nuclear DNA analysis has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between an evidentiary sample and a specific individual or source.” Professor Berger served on the Committee that wrote the report.

The centerpiece of the report is a recommendation that Congress establish an independent federal entity, the National Institute of Forensic Sciences (NIFS), which would, among other things, fund research “to address issues of accuracy, reliability, and validity in the forensic science disciplines” and establish and enforce “best practices” for forensic science professionals and laboratories. Other recommendations include (1) mandating laboratory accreditation and practitioner certification, (2) removing crime laboratories from administrative control of law enforcement agencies and prosecutors’ offices, (3) supporting investigations into human

---

76 Margaret A. Berger, Laboratory Error Seen Through the Lens of Science and Policy, 30 U.C. DAVIS L. REV. 1081 (1997).
80 Id. at 100.
81 Id. at 22-23 (Recommendation 3.).
82 Id. at 19 (Recommendation 1(a)).
83 Id. (Recommendation 1(b)).
84 Id. at 24 (Recommendation 4.).
observer bias and sources of human error in forensic analysis, and (4) developing standard terminology and model laboratory report formats.

According to the Report, NIFS should also draft a code of ethics for all forensic sciences and encourage individual forensic societies to incorporate this national code as part of their professional codes of conduct.” In 1982, Professor Berger raised this issue at a conference on scientific evidence: “Is there, or should there be, a code of conduct for scientists that would provide a basis for objecting to testimony that strayed too far from strict impartiality?” At the same time, she commented that “[t]he inequality of available resources between the prosecution and the defense is alarming enough, but it is compounded by the lack of scientific literacy on the part of most defense lawyers.”

Recommendation 10 of the NAS Report provides, in part: “NIFS should also support law school administrators and judicial education organizations in establishing continuing legal education programs [on forensic science] for law students, practitioners, and judges.”

CONCLUSION

Professor Berger is one of the few scholars who is equally comfortable examining expert testimony issues in both civil and criminal cases. Her contribution to the development
of the law and society, through her scholarship and public service on various NAS committees, has been exemplary. It is only fitting that this essay should conclude with her words:

What criminal defendants need in order to deal more effectively with the forensic identification expertise proffered against them is not more Daubert, but tools that would enable them to make more cogent evidentiary arguments—better counsel, access to expert assistance and more discovery.  

Berger, supra note 75, at 1140. She also wrote:

I strongly believe that we need a very stringent standard of proof in criminal cases. I do not think, however, that Daubert v. Merrell Dow Pharmaceuticals, Inc. has been productive in effectuating this goal. In civil cases, courts engage in rigorous gatekeeping and often exclude plaintiffs' experts because the theory underlying their testimony has not been adequately validated. But I see no sign of a parallel approach in criminal cases even when there are problems with the assumptions on which the prosecution's expert testimony rests.

Id. at 1125.
Professor Margaret Berger, the Epitome of the Fully Engaged Scholar and Friend of the Court

Edward J. Imwinkelried†

Today’s law professors are no longer content to remain in the “ivory tower.” Rather, they aspire to be fully engaged in the process of law reform. They not only hope that their scholarship will be creative and theoretically sound; they also want it to have real world impact. Margaret Berger’s scholarly career is a model for any academic who entertains that aspiration.

To be sure, Professor Berger is a prominent figure within the “ivory tower.” For decades, she has been a coauthor of one of the leading evidence casebooks, *Evidence: Cases and Materials*, with Judge Weinstein and Professors Mansfield and Abrams.1 Moreover, she has published widely cited articles in many of the most highly regarded law reviews.2

However, her influence extends far beyond the world of legal education. Law reform organizations have often turned to her for guidance and insight. She has been a member of several National Academy of Sciences committees, including the Committee on DNA Technology in Forensic Science. She was the Reporter for the Post-Conviction Issues Working Group of the National Commission on the Future of DNA Evidence. The

---

1 Edward L. Barrett, Jr. Professor of Law, University of California Davis.
prestigious Carnegie Commission on Science, Technology, and Government has also called on her as a consultant.

For its part, the practicing bar pays special attention to Professor Berger’s writing. She is the coauthor of the foremost treatise on federal evidence law, *Weinstein’s Federal Evidence: Commentary on Rules of Evidence for the United States Courts and Magistrates*. I teach Trial Practice. I devote one class session to evidentiary objections. In that session, I discuss the question of which secondary authorities the trial attorney should cite to the judge. I have told literally thousands of students that when you have time at sidebar to cite only one authority to a federal judge, that authority should be *Weinstein’s Evidence*. It undeniably carries more weight with sitting federal District Court judges than any other treatise or text.

Judges not only have a high regard for Professor Berger’s contributions to the Weinstein treatise; she has published other works that are typically at the fingertips of federal judges. After the Supreme Court handed down its decision in *Daubert*, the Federal Judicial Center decided that it needed to provide the federal judiciary with research tools to help judges deal more knowledgeably with scientific issues. The center has released two editions of its celebrated *Reference Manual on Scientific Evidence*. Every federal District Court judge in the United States has that text either on the bench or in chambers. Professor Berger contributed substantial articles to both editions. In the first edition, she authored “Evidentiary Framework,” which gave judges an overview of the impact of *Daubert*. The second edition includes her article, “The Supreme Court’s Trilogy on the Admissibility of Expert Testimony.” That article not only contains further perspective on the original *Daubert* decision; the article adds a discussion of *Daubert*’s progeny, *Joiner* and *Kumho*. It was expectable that

---

when the Judicial Conference reconstituted the Advisory Committee on the Federal Rules of Evidence, Professor Berger was named the first Reporter for the committee.

These facts give a sense of the extent of Professor Berger’s influence. However, in this article I would like to focus on the considerable influence she has had as a friend of the Court—as the author of amicus briefs submitted to the United States Supreme Court. During her long career, Professor Berger has submitted a large number of amicus briefs in cases pending before the Court. However, two amicus briefs are especially noteworthy, namely, her amicus briefs in the original Daubert litigation and her brief in the subsequent Kumho case. A careful comparison of the contents of Professor Berger’s briefs in those cases and the Court’s ultimate opinions reveals the remarkable degree to which Professor Berger’s arguments seemingly influenced the Court’s ruling and reasoning in both decisions.

I. Professor Berger’s Amicus Brief in Daubert

In 1992, Professor Berger was the lead author of an amicus brief in Daubert on behalf of the Carnegie Commission on Science, Technology, and Government. In the long term, one of the most important passages in Daubert will prove to be Justice Blackmun’s observation that “arguably, there are no certainties in science.” Prior to Daubert, many courts had

---

13 Daubert Amicus Brief, supra note 11.
subscribed to the naïve belief that at least the “exact” sciences could yield absolutely certain conclusions. Justice Blackmun shattered that naïveté in Daubert. In her amicus brief in Daubert, Professor Berger had urged the Court to do precisely that. She noted that in the past, the courts had “assume[d] that there is much more definiteness in science than actually exists.” As she described the scientific process, even in fields such as physics and chemistry the experimental/observational methodology yields only “contingent,” “provisional” conclusions. Since it is always conceivable that a subsequent experiment will falsify a hypothesis supported by earlier experiments, scientific investigators cannot lay claim to “final or permanent” truth.

In large part, the Court agreed with Professor Berger’s position because the Court embraced her conception of the scientific process itself. The brief repeatedly described the essence of the scientific method as the “formulati[on] [of] hypotheses” and “[r]igor[ous] . . . testing of [the] hypotheses” to validate or falsify them. Justice Blackmun’s description of the scientific method in his lead opinion in Daubert is strikingly similar: “a process for proposing and refining theoretical explanations about the world that are subject to further testing and refinement.”

The Daubert Court drew a number of doctrinal implications from its conclusions about the nature of the scientific enterprise—the very implications that Professor Berger identified in her amicus brief. First, Justice Blackmun abandoned the traditional general acceptance test for the admissibility of scientific testimony. The Justice characterized

---

16 Daubert Amicus Brief, supra note 11, at *9.
17 Id. at *5.
18 Id. at *9.
19 Id. at *4.
20 Id.
21 Id.
22 See id. (“how science is conducted”); id. at *8 (“an extensive examination of the hypotheses being put forth”); id. at *10 (“the process by which the theory was generated or tested”).
24 See id. at 587-89.
the test as too “austere [a] standard.”\textsuperscript{25} He criticized the test as being “at odds with the ‘liberal thrust’ of the Federal Rules [of Evidence].”\textsuperscript{26} In her brief, Professor Berger had asserted that a realistic understanding of the nature of scientific methodology “require[d] rejection” of the 1923 \textit{Frye}\textsuperscript{27} case that announced the general acceptance test.\textsuperscript{28} She explained that the traditional test was “simplistic”\textsuperscript{29} and “incompatible with the essence of the scientific endeavor.”\textsuperscript{30} She wrote that the \textit{Frye} test, if “taken literally[,] rejects valuable insights that bear all the hallmarks of acceptable science.”\textsuperscript{31}

Justice Blackmun supplanted the traditional standard with an essentially methodological test.\textsuperscript{32} He declared that the focus should be on the soundness of the scientific methodology supporting the expert’s opinion\textsuperscript{33} rather than the judge’s view of the correctness of the conclusion reached by the expert. The Justice elaborated that the opinion’s proponent must convince the trial judge that the opinion is “derived by the scientific method,”\textsuperscript{34} that is, “supported by appropriate validation.”\textsuperscript{35} Those passages echoed the part of Professor Berger’s amicus brief in which she argued that

[t]he question is not whether the judge agrees with the results of the study . . . . Rather the court must decide whether the study was set up and carried out in a manner that conforms to standards in the scientific community.\textsuperscript{36}

Professor Berger’s brief even anticipated the manner in which Justice Blackmun would rationalize his holding as a matter of statutory interpretation. He reasoned that “[t]he primary locus” for deriving the test was Rule 702.\textsuperscript{37} More specifically, he ruled that when marshaling testimony about

\begin{flushleft}
\textsuperscript{25} Id. at 589.

\textsuperscript{26} Id. at 588 (citation omitted) (internal quotation marks omitted).

\textsuperscript{27} See \textit{Frye} v. United States, 293 F. 1013 (D.C. Cir. 1923).

\textsuperscript{28} \textit{Daubert} Amicus Brief, supra note 11, at *2, *7.

\textsuperscript{29} Id. at 7.

\textsuperscript{30} Id.

\textsuperscript{31} Id.


\textsuperscript{34} Id. at 590.

\textsuperscript{35} Id.

\textsuperscript{36} \textit{Daubert} Amicus Brief, supra note 11, at *17; see also id. at *12 (“the methodology”).

\textsuperscript{37} \textit{Daubert}, 509 U.S. at 589.
\end{flushleft}
the scientific methodology underpinning the expert’s opinion, the proponent must demonstrate the expert’s reasoning is reliable “scientific . . . knowledge” within the meaning of that expression in Rule 702.38 This was the identical statutory basis that Professor Berger’s brief singled out. Citing Rule 702, she stated that the trial judge ought to inquire whether the expert’s reasoning “conform[s] to the characteristics of ‘scientific knowledge.’”39

After announcing the general methodological test, Justice Blackmun proceeded to list several factors that trial judges should consider in deciding whether the expert’s opinion rests on sound scientific methodology.40 His list is quite similar to the list of such factors included in Professor Berger’s amicus brief. The Justice’s list includes these factors:

- Whether the hypothesis is empirically testable;41
- Whether it has been tested;42
- Whether the research has been subjected to peer review (although he cautioned that peer review “is not a sine qua non of admissibility”);43
- Whether the hypothesis is generally accepted to the extent that general acceptance is circumstantial evidence that other scientists have scrutinized the research and found it to be methodologically sound;44 and
- Whether the expert’s methodology has a known or ascertainable error rate.45

The list in Professor Berger’s amicus is remarkably parallel:
- Whether the theory “is capable of being proven false through observation or experimentation”;46
- Whether the theory “has in fact been subjected to an empirical scrutiny”,47

---

38 Id. at 589-90 (quoting FED. R. EVID. 702).
39 Daubert Amicus Brief, supra note 11, at 11-12.
40 Daubert, 509 U.S. at 593-94.
41 Id. at 593.
42 Id.
43 Id.
44 Id. at 594.
45 Id.
46 Daubert Amicus Brief, supra note 11, at *13.
47 Id. at *14.
Whether the research has been peer reviewed (although she cautioned that peer review should not be an invariable requirement); 48
Whether the theory has been generally accepted to the extent that such acceptance is circumstantial proof that other scientists have concluded that the underlying research was “produced in conformity with the scientific process”; 49 and
What the technique’s “error rate” is. 50

Professor Berger’s brief not only sketched the basic outline of Justice Blackmun’s opinion, it also furnished some of the fine print. After describing the validation test he derived from Rule 702, the Justice cited Rule 401 and added that to be relevant, the expert’s theory must “fit” the specific facts of the case. 51 In her amicus brief, Professor Berger cited the same statute and emphasized that to satisfy Rule 401, the expert’s research has to “fit” the “facts in the case.” 52 Procedurally, Justice Blackmun stressed that Federal Rule 104(a) governs the trial judge’s determinations under the validation test. 53 Professor Berger made precisely that point in her brief. 54 In short, to a considerable extent, Professor Berger’s amicus brief presaged the content of Justice Blackmun’s opinion in Daubert. The coincidence is so extensive that the conclusion is well nigh unavoidable that her brief was a major influence on the Daubert Court’s decision.

II. PROFESSOR BERGER’S AMICUS BRIEF IN KUMHO

Six years after its Daubert decision, the Supreme Court revisited the topic of expert testimony in Kumho Tire Co., Ltd. v. Carmichael. 55 Kumho raised the question of the standard for determining the admissibility of non-scientific expert testimony. 56 As in Daubert, Professor Berger was the lead

48 Id. at *25-28.
49 Id. at *5-6.
50 Id. at *16-17.
52 Daubert Amicus Brief, supra note 11, at *11, n.11.
53 Daubert, 509 U.S. at 592.
54 Daubert Amicus Brief, supra note 11, at *12.
56 Id. at 141.
author of an amicus brief in *Kumho.* Just as Justice Blackmun’s opinion in *Daubert* reflected the persuasiveness of her brief in that case, it is easy to discern the imprint of Professor Berger’s amicus brief on Justice Breyer’s opinion in *Kumho.* Professor Berger’s brief urged a balanced approach which Justice Breyer endorsed in his opinion.

On the one hand, in her brief Professor Berger argued that the rigorous validation standard enunciated in *Daubert* was sometimes inappropriate for assessing the reliability of non-scientific expertise such as medical testimony. In his opinion, Justice Breyer concurred, observing that *Daubert* “referred only to ‘scientific’ knowledge” because “that [wa]s the nature of the expertise’ at issue” there. Elaborating, Professor Berger asserted that in a case involving non-scientific expertise, it would sometimes be wrong-minded to apply the factors enumerated in *Daubert.* Justice Breyer approved of that view in his opinion.

Next, Professor Berger generally cautioned against attempting to “construct a complex,” rigid classification system of types of expertise. The Justice agreed, stating that “it would prove difficult, if not impossible, for judges to administer evidentiary rules under which a gatekeeping obligation depended upon . . . distinction[s]” among various kinds of expert testimony. In his judgment, “conceptual efforts” to fashion such sharp distinctions were “unlikely to produce clear . . . lines capable of application in particular cases.” More specifically, Professor Berger asserted that, in at least some

---

57 *Kumho Amicus Brief, supra note 12.*
58 See id. at *4, *18.
60 See *Kumho Amicus Brief, supra* note 12, at *4.
61 *Kumho Tire Co.*, 526 U.S. at 150-51 (“We agree with the Solicitor General that ‘[t]he factors identified in *Daubert* may or may not be pertinent in assessing reliability, depending on the nature of the issue, the expert’s particular expertise, and the subject of his testimony.’. . . [W]e can neither rule out, nor rule in, for all cases and for all time the applicability of the factors mentioned in *Daubert,* nor can we now do so for subsets of cases categorized by category of expert or by kind of evidence” (first alteration in original) (citation omitted) (quoting Brief for the United States as Amicus Curiae Supporting Petitioners at 19, *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999) (No. 97-1709), 2002 WL 541947).
63 *Kumho Tire Co.*, 526 U.S. at 148.
64 Id.; see also id. at 151 (“We do not believe that Rule 702 creates a schematism that segregates expertise by type while mapping certain kinds of questions to certain kinds of experts.”).
instances, an expert’s experience could be an adequate foundation for an opinion.\textsuperscript{65} Justice Breyer also posited that a witness’s experience can be a sufficient basis for an expert opinion.\textsuperscript{66}

On the other hand, Professor Berger argued that the trial judge must demand that the expert’s proponent demonstrate that the expert’s opinion amounts to more than the witness’s “subjective belief.”\textsuperscript{67} For his part, Justice Breyer came to the same conclusion.\textsuperscript{68} In her brief, Professor Berger asserted that “experience-based knowledge should not be automatically inadmissible . . . .”\textsuperscript{69} Likewise, Justice Breyer stressed that the trial judge must scrutinize even “experienced-based testimony.”\textsuperscript{70} Professor Berger contended that the judge ought to insist that in preparing his or her testimony, the expert “exercis[ed] the same level of intellectual rigor that generally characterizes that expert’s field of expertise.”\textsuperscript{71} In formulating his holding, Justice Breyer echoed Professor Berger’s brief; he wrote that the trial judge must ensure that “an expert, whether basing testimony on professional studies or personal experience, employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.”\textsuperscript{72}

In the final analysis, Professor Berger called on the Court to grant trial judges discretion both in applying Rule 702’s substantive admissibility standard and in devising procedures for doing so. Substantively, Professor Berger recommended that the Court “accord the trial judge a substantial measure of discretion” in selecting the factors to be used in gauging the reliability of nonscientific expertise.\textsuperscript{73} She counseled against “bright-line test[s].”\textsuperscript{74} Procedurally, while

\textsuperscript{65} See Kumho Amicus Brief, supra note 12, at *6-7.
\textsuperscript{66} See Kumho Tire Co., 526 U.S. at 152.
\textsuperscript{67} Kumho Amicus Brief, supra note 12, at *13 (quoting Daubert v. Merrell Dow Pharm., 509 U.S. 579, 590 (1993))).
\textsuperscript{68} See Kumho Tire Co., 526 U.S. at 147-48 (“In Daubert, the Court specified that it is the Rule’s word ‘knowledge’ . . . that ‘establishes a standard of evidentiary reliability.’ Hence, as a matter of language, the Rule applies its reliability standard to all ‘scientific,’ ‘technical, or ‘other specialized’ matters within its scope.” (quoting Daubert v. Merrell Dow Pharm., 509 U.S. 579, 589-90 & n.8 (1993))).
\textsuperscript{69} Kumho Amicus Brief, supra note 12, at *19.
\textsuperscript{70} Kumho Tire Co., 526 U.S. at 151.
\textsuperscript{71} Kumho Amicus Brief, supra note 12, at *3.
\textsuperscript{72} Kumho Tire Co., 526 U.S. at 152.
\textsuperscript{73} Kumho Amicus Brief, supra note 12, at *14.
\textsuperscript{74} Id. at *3.
some had asked the Court to recognize a right to a pretrial *Daubert* hearing, in her brief Professor Berger staked out the position that the trial judge should also have a significant measure of discretion in fashioning procedures for administering Rule 702. In his opinion, Justice Breyer came down on both issues in the same fashion. On the substantive question, the Justice stated that “in a particular case,” the trial judge has “broad latitude” in choosing the factors that are “reasonable measures of reliability.” Procedurally, the Justice declared:

The trial court must have the same kind of latitude in deciding how to test an expert’s reliability, and to decide whether or when special briefing or other proceedings are needed to investigate reliability . . . . That standard applies as much to the trial court’s decisions about how to determine reliability as to its ultimate conclusion.

**CONCLUSION**

It is fair to say that *Daubert* and *Kumho* are the two most important expert testimony opinions ever rendered by the United States Supreme Court. They not only control in federal court; they have been widely cited and followed by state courts as well. A comparison of those decisions with Professor Berger’s amicus briefs reveals a remarkable degree of similarity between the views she urged and the positions ultimately taken by the Court. At the very least, Professor Berger is an incredible prognosticator. More likely, though, her briefs were instrumental in convincing the Court to embrace those positions. By venturing beyond the “ivory tower” and joining the fray in *Daubert* and *Kumho*, Professor Berger helped shape two of the most important evidence decisions of this era. To a degree, her amicus briefs provided the Court with blueprints for those decisions, just as her distinguished career has become the blueprint for any member of the academy who aspires to take up the challenge of engaging in real world law reform.

---

75 See id. at *21-26; see also Berger, Procedural Paradigms for Applying the Daubert Test, supra note 2, at 1361-63.
76 *Kumho Tire Co.*, 526 U.S. at 153; see also id. at 152 (“[A] trial court should consider the specific factors identified in *Daubert* where they are reasonable measures of the reliability of expert testimony.”).
77 Id. at 152 (emphasis omitted).
Probability, Individualization, and Uniqueness in Forensic Science Evidence

LISTENING TO THE ACADEMIES

David H. Kaye

INTRODUCTION

These are dark days for the forensic sciences. Newspaper and magazine articles, op-ed headlines, television news, and radio talk shows refer to the field as “seriously deficient,” a “dismal science,” “in disarray,” even “clueless.” The stimulus for this negative publicity blitz is a report of a congressionally-mandated “independent forensic science committee at the National Academy of Sciences.” This belated report calls for structural and cultural changes in the forensic science community ranging from the separation of laboratories and police departments, to a uniform, enforceable code of ethics and standardized testimony. The report found that:

© 2009 D.H. Kaye. All rights reserved.
† Distinguished Professor of Law, Weiss Family Scholar, and Graduate Faculty, Forensic Science Program, Pennsylvania State University. I am grateful to Jay Koehler and Stephen Stigler for comments on a draft of this paper.
6 The report originally was scheduled for January 2008. See id. (announcing a starting date of September 2006 for a project of 16 months duration).
The forensic science disciplines exhibit wide variability with regard to techniques, methodologies, reliability, level of error, research, general acceptability, and published material. . . . Many of the processes used in the forensic science disciplines are . . . not based on a body of knowledge that recognizes the underlying limitations of the scientific principles and methodologies for problem solving and discovery. . . . Some of these activities [encompassed by the term “forensic science”] might not have a well developed research base, are not informed by scientific knowledge, or are not developed within the culture of science.9

These observations are hardly news to the other academy—the professoriate. For years, the authors of legal treatises and journals have complained bitterly about the lack of regulation of forensic laboratories, the absence of rigorous proficiency testing, and the dearth of basic research that would demonstrate the alleged ability of fingerprint, toolmark, and other analysts to identify traces from one person or object to the exclusion of all others in the world.10 They have written dismissively of “nonscience forensic sciences” that “have little or no basis in actual science,” and they have implored courts to exclude testimony, pending better research showing that analysts can live up to their claims.11 With rare exceptions, however, the courts have failed to perceive the gap between

---

9 Id. at 38-39.


11 E.g., Simon A. Cole, Does “Yes” Really Mean Yes? The Attempt to Close Debate on the Admissibility of Fingerprint Testimony, 45 JURIMETRICS J. 449 (2005); Lyn Haber & Ralph Norman Haber, Experiential or Scientific Expertise, 7 L., PROBABILITY & RISK 143 (2008); D. Michael Risinger & Michael J. Saks, Science and Nonscience in the Courts: Daubert Meets Handwriting Identification Expertise, 82 IOWA L. REV. 21 (1996); Adina Schwartz, A Systemic Challenge to the Reliability and Admissibility of Firearms and Toolmark Identification, 6 COLUM. SCI. & TECH. L. REV. 1, 1 (2005) (“[A]ll firearms and toolmark identifications should be excluded until adequate statistical empirical foundations and proficiency testing are developed for the field.”).
optimistic theory and hard proof, and they have accepted remarkably weak forms of validation.\textsuperscript{12}

With the imprimatur of the National Academy of Sciences behind recommendations for major change, the need for forensic scientists or analysts to retreat from the most extreme claims finally should be apparent to the judiciary as well as the forensic science community. But how far should this retreat go? Should forensic scientists be forever barred from giving an opinion that a DNA sample, a fingerprint, or a broken part of an object originated from a particular person, finger, or matching object? What does it take to justify such opinions? This essay seeks to clarify these questions by scrutinizing several statements and recommendations on how to present testimony offered in an essay by two of the legal academy’s foremost critics of contemporary forensic science. In \textit{The Individualization Fallacy in Forensic Science ("Fallacy")},\textsuperscript{13} Professors Michael Saks and Jay Koehler make the following statements:

(1) The concept of “individualization,” which lies at the core of numerous forensic science subfields, exists only in a metaphysical or rhetorical sense. It has no scientific validity, and it is sustained largely by the faulty logic that equates infrequency with uniqueness.\textsuperscript{14}

(2) Application of the product rule necessarily falls short of establishing unique individualization. The product of probabilities greater than zero always yields a value greater than zero. The probabilistic approach, therefore, always leads to the conclusion that a source other than the suspected individual or object might exist.\textsuperscript{15}

(3) The claim of unique individuality cannot be proven with samples . . . . “It is impossible to prove any human characteristic to be distinct in each individual without checking every individual . . . .” Anything less results in probability statements rather than conclusions of absolute specificity and absolute identification.\textsuperscript{16}

\textsuperscript{12} See, e.g., D.H. Kaye et al., supra note 10; 4 Modern Scientific Evidence: The Law and Science of Expert Testimony, supra note 10, at 276-77; Mnookin, supra note 10, at 127-29.


\textsuperscript{14} Id. at 205.

\textsuperscript{15} Id. at 209 (footnote omitted).

\textsuperscript{16} Id. at 211 (footnote omitted). After asserting that sampling is incapable of proving uniqueness, the first sentence adds that this is “especially” so for “samples that are a tiny proportion of the relevant population.” Id.
(4) As Karl Popper famously explained, it is logically impossible to prove a hypothesis by accumulating positive instances. The hypothesis, “all swans are white,” remains unproven, even after a large number of sightings of white swans, because the sighting of a single black swan would disprove it. Similarly, the hypothesis that no two objects are indistinguishably alike cannot be proven true from an accumulation of observations in which different object sources produce distinctive markings.\(^\text{17}\)

(5) Even a very large number of pairwise, case-by-case comparisons made by individual examiners would not provide a satisfactory method for testing the object uniqueness claim.\(^\text{18}\)

Although I agree with the critique of a great deal of forensic science testimony, to the extent that these statements imply that, even in principle, science cannot establish the uniqueness of objects, I am dubious. In addition, individualization—the conclusion that “this trace came from this individual or this object”—is not the same as, and need not depend on, the belief in universal uniqueness.\(^\text{19}\) Consequently, there are circumstances in which an analyst reasonably can testify to having determined the source of an object, whether or not uniqueness is demonstrable. Part I of this essay shows why the arguments for radical skepticism of uniqueness are not convincing. Part II explains the distinction between individualization and uniqueness. It explicates what I believe to be the real individualization fallacy—the putatively sharp dichotomy between class and individual characteristics. Part III applies these ideas to courtroom testimony and argues that a variety of courtroom explanations of the meaning of a match should be permissible.

I. PROVING UNIQUENESS

A. Metaphysics

If all that a criminalist can say is that, in some untestable ways, no two objects are the same, then the testimony should be excluded as irrelevant and as not constituting specialized “knowledge” within the meaning of

\(^\text{17}\) Id. at 212 (footnotes omitted).

\(^\text{18}\) Id.

Rule 702 of the Federal or Uniform Rules of Evidence.\textsuperscript{20} Courtroom claims of individuality, however, necessarily have to do with measurable characteristics that can exhibit unequivocal differences and similarities.\textsuperscript{21} For example, in 1992, a committee of the National Academy of Sciences (the “NAS”) recommended that “[c]ourts should take judicial notice of [the] scientific underpinnings of DNA typing”—including the fact that “[e]ach person’s DNA is unique (except that of identical twins) . . . .”\textsuperscript{22} Although we lack the technology to generate error-free sequences of the more than six billion base pairs that constitute diploid human genomes, in principle, the claim of individuality can be refuted by a much improved sequencing experiment that establishes perfect congruence in two individual genomes. The NAS committee’s individualization hypothesis—which is standard fare in human genetics—could be wrong, but it is not metaphysical.\textsuperscript{23}

B. The Product Rule and Nonzero Probabilities

Saks and Koehler’s second point is that uniqueness is beyond the realm of proof because “[t]he product of probabilities greater than zero always yields a value greater than zero. The probabilistic approach, therefore, always leads to the conclusion that a source other than the suspected individual or object might exist.”\textsuperscript{24}

This argument proves too much. If the problem is simply that another source \textit{might} exist, then the fallacy infects all scientific research and testimony. There is always some nonzero probability of an erroneous conclusion. Ohm’s law


\textsuperscript{21} See, e.g., Michael G. Koot et al., Radiographic Human Identification Using Bones of the Hand: A Validation Study, 50 J. FORENSIC SCI. 263, 263-64 (2005) (suggesting that “skeletal features formed in late childhood remain unique throughout life” on the basis, in part, of a previous analysis of 40 pairs of hand radiographs of same-sex, identical twins for which “there were, in every instance, some features which made it possible to distinguish the hand and wrist bones of one person from those of his or her own twin”). This limited study falls woefully short of demonstrating uniqueness, but the logic of looking at identical twins for differences is sound.

\textsuperscript{22} NAT’L RESEARCH COUNCIL COMM. ON DNA TECH. IN FORENSIC SCI., DNA TECHNOLOGY IN FORENSIC SCIENCE 23 (1992) [hereinafter NRC 1992].

\textsuperscript{23} In describing certain claims of uniqueness as “not metaphysical,” I am responding to “the impression that metaphysics is a study that somehow ‘goes beyond’ physics.” Peter van Inwagen, \textit{Metaphysics}, in \textit{STANFORD ENCYCLOPEDIA OF PHILOSOPHY} (2007), http://plato.stanford.edu/entries/metaphysics/.

\textsuperscript{24} Saks & Koehler, supra note 13, at 209.
might not be exactly right, or it might break down tomorrow, but electrical engineers can safely assume that it is absolutely true. Returning to forensic science, was it fallacious for the 1992 NAS committee to represent that an individual’s full genome is unique? A researcher applying standard statistical reasoning would reject the hypothesis of duplication vis-à-vis the alternative hypothesis of uniqueness when the probability of duplication in the population is small enough. A second NAS committee suggested that “[w]ith an increasing number of loci available for forensic analysis, we are approaching the time when each person’s profile will be unique (except for identical twins and other close relatives).” Its 1996 report distinguished between specific and general claims of uniqueness. A specific profile might be unique: “Suppose that, in a population of $N$ unrelated persons, a given DNA profile has probability $P$. The probability (before a suspect has been profiled) that the particular profile observed in the evidence sample is not unique is at most $NP$. A small probability $NP$ indicates that the one profile under consideration is likely to be unique within a population that contains as many as $N$ unrelated people. This is uniqueness conditioned on a given genotype.

General uniqueness refers to all the profiles in the population. “A lower bound on the probability that every person is unique depends on the population size, the number of loci, and the heterozygosity of the individual loci.” With some simplifying assumptions, the probability of this event also can be estimated. “Neglecting population structure and close relatives, 10 loci with a geometric mean heterozygosity of 95%...”

In their rejoinder in this issue of the Law Review, Professors Koehler and Saks question whether Ohm’s law (as opposed to say, the subsequent example of the laws of electromagnetism) is a good illustration of the common practice of expressing textbook knowledge as established scientific fact. Jonathan J. Koehler & Michael J. Saks, Individualization Claims in Forensic Science: Still Unwarranted, 75 BROOK. L. REV. 1187, 1203-04 (2010). They are correct in criticizing this first example. More dubious, however, is their insistence that “[t]he implication is that even if forensic examiners can’t be 100% sure of their ability to individualize, they are safe in proceeding on the assumption that their individualization conclusions are absolutely true.” Id. at 1203. Whether it is safe to accept any scientific statement obviously depends on its foundation. See infra note 45. As the next sentence in the text and the larger discussion plainly indicate, the only implication of these examples is that, in appropriate circumstances, statements of uniqueness (such as the textbook claims about human genomes) could be accepted as true despite a nonzero probability that they are false.
give a probability greater than about 0.999 that no two unrelated people in the world have the same profile.”

For Saks and Koehler, however, no probability of duplication is small enough to warrant an opinion that DNA or anything else is unique. Thus, they reject the reasoning that a “probability of two individuals having the same fingerprint is one out of $1 \times 10^{60}$ . . . is so small as to exclude the possibility of any two individuals having the same fingerprints.” They are correct, but only in the trivial sense that every event with a nonzero probability is a “possibility.” $P = 10^{-60}$ is supposed to be the probability that two randomly selected people will have matching fingerprints. Although I doubt the accuracy of the estimated match probability, the allegedly “faulty logic”—the move from $P = 10^{-60}$ for the probability of a match to a randomly selected pair to zero for the probability of a match for all possible pairs—is defensible. Suppose that the world’s population ($N$) is seven billion. The number of distinct pairs of people is $N(N - 1)/2$, which is on the order of $10^{19}$. Even for this many comparisons, when each has only a probability of $10^{-60}$ of being the same, the chance of one or more identical fingerprints in the world’s population is about $10^{-41}$. Technically, this

---

29 Id. The computation, “an application of the ‘birthday problem’ with unequal probabilities,” can be found in appendix 5C of the report. More STR loci (the type currently used in DNA identification) would be required to achieve the same probability of uniqueness. See id. at 165; see also KAYE ET AL., supra note 10, § 12.5.3.

30 Saks & Koehler, supra note 13, at 203 (quoting RICHARD SAFERSTEIN, CRIMINALISTICS: AN INTRODUCTION TO FORENSIC SCIENCE 73 (9th ed. 2007)).

31 If we truly could believe that the chance of the fingerprints of any two different people matching were $10^{-60}$, we could—and should—believe in the uniqueness of fingerprints. Whether fingerprint examiners reliably can differentiate latent prints at the level of detail needed to give rise to probabilities such as $10^{-60}$ is obviously another story (and not a very believable one). See generally David H. Kaye, From Snowflakes to Fingerprints: A Dubious Courtroom Proof of the Uniqueness of Fingerprints, 71 INT’L STAT. REV. 521, 524 (2003) (criticizing an unpublished study by an FBI contractor introduced by the Department of Justice and relied on by federal courts to show “that the probability of finding two people with identical fingerprints was one in ten to the ninety-seventh power [and] that the probability of finding two different, partial fingerprints to be identical was one in ten to the twenty-seventh power.”); see also Christophe Champod, Fingerprint Examination: Towards More Transparency, 7 L., PROBABILITY & RISK 111, 113 (2008) (“Systematic research on the selectivity of fingerprint features [indicates] that even very limited configurations of fingerprint minutiae can provide . . . match probabilities on the order of 1 in a billion, even without considering the statistical contribution of level 1 features (general pattern, ridge counts, etc.) or other fingerprint features if available.”).

32 Saks & Koehler, supra note 13, at 204.

33 Intuitively, if each comparison has the same tiny probability $P$ of producing a match and we have made $n$ comparisons, then the probability of at least one match is close to $nP$. See Frederick Mosteller, Understanding the Birthday Problem, 55 MATH. TEACHER 322 (1962) (eqn. 7), reprinted in SELECTED PAPERS OF FREDERICK MOSTELLER
probability is greater than zero, but that mathematical truism
hardly makes it fallacious to exclude as totally unrealistic the
thought of a matching fingerprint from someone else. It is not a
fallacy to infer uniqueness (both specific and general) when the
match probability $P$ is immensely smaller than the reciprocal of
the size of a population of objects, every one of whose members
has the small probability $P$ of matching.34

Thus, the problem with using probability theory to
demonstrate uniqueness is not that the probability of
duplication always exceeds zero. The difference might be too
small to matter. Such demonstrations are generally
unconvincing because it is so hard to establish that the models
are sufficiently realistic and accurate to trust the computed
probabilities. But sometimes probabilities are negligible. Just
think about the chance that you would suffocate because all the
nearby molecules of oxygen in the room would happen to move
to the other half of the room. A few simple assumptions and a
bit of statistical mechanics demonstrate that the possibility
need not worry us.

C. Direct Testing (Sampling)

Saks and Koehler’s remaining arguments (3 through 5)
boil down to the claim that sample data cannot establish the
exact proportion of an entire population that shares a given
characteristic. Uniqueness means that the proportion of objects
with the given feature in the whole population of size $N$
is exactly $1/N$. Yet, no matter how close the sample proportion
comes to $1/N$, the next sample datum could establish that the
population proportion is $2/N$ or more.

Again, this is true but not indicative of faulty reasoning.
Certainly, there are reasons to distrust the “it hasn’t happened
yet” theory of uniqueness. Fallacy cogently explains the
limitations of unsystematic, “pairwise, case-by-case
comparisons made by individual examiners,”35 and it points to

349, 351 (Stephen E. Fienberg & David Caster Hoaglin eds., 2006). For $n = 10^{19}$ and $P = 10^{-60}$, this gives $10^{-41}$. (A more precise calculation gives a value with the same order of magnitude.) Since the probability of the world’s population having distinct fingerprints
differs from 1 by a mere $10^{-41}$, for all practical (and impractical) purposes, fingerprints
are globally unique—that is, if people are assigned fingerprint patterns by a process
that is equivalent to randomly sampling an infinite population in which the $10^{60}$
distinct patterns are uniformly distributed.


35 Id.
contradictory reports of matches in trace evidence coming from different individuals. These are reasons enough for skepticism, but the white swans and the insistence that only a census will do the job makes it look as if the examiners are pursuing a line of proof that is logically incapable of supporting the desired inference. The flaw is not with the logic. It is with the data. The number of comparisons required to prove uniqueness by brute empirical force (that is, to make the normal statistical inference from a sample to a population) is surprisingly larger than one might think. It takes a random-match probability whose reciprocal is orders of magnitude larger than the population of objects to make uniqueness almost certain.

But this difficulty with a direct empirical proof is not tantamount to Popper’s realization that universal laws cannot be proved to a logical certainty by simple induction. Modern science is full of universal laws. The laws of electromagnetism, for instance, remain unproven and unprovable in Popper’s logical sense. The only universal propositions that can be proven to a certainty are deductively valid ones, such as the theorem that all whole numbers that end in an even digit are divisible by two. No experimentation is required to test this law. In contrast, no matter how many times scientists observe that a change in the magnetic flux in a coil or wire induces a voltage, they cannot be certain that it will happen the next time. In principle, a single experiment with no change in voltage would disprove Faraday’s Law.

In Popper’s framework, the repeated failure of forensic examiners to find “two sets of markings produced by different

---

36 Id. at 213.

37 Fallacy can be read this way as well. See id. (“As the size of a comparison database becomes larger, the object uniqueness hypothesis is subjected to an increasingly tough empirical test. If, under these circumstances, scientists still do not find indistinguishably similar matches produced by different objects, then object uniqueness becomes a more credible theory.”).

38 Id. at 203-04 (describing a version of the “birthday problem” in probability theory).

39 Id. at 148.

40 Saks & Koehler, supra note 13, at 212.

41 See id.

42 Id. In practice, a single experiment would not suffice. Effects that cannot be replicated are likely to be the result of experimental error. Because data gathering is fallible, a single sighting of a black swan might not be enough to disprove the hypothesis that all swans are white. The observer might have been mistaken about the bird (it was not really a swan) or its true color (it had flown through a cloud of soot).
sources that are indistinguishable from each other\textsuperscript{43} lends some support to the generalization that all different sources produce distinguishable markings. His swan example merely shows that we never can be absolutely certain of any generalization. If that is all that the individualization fallacy consists of, then all induction is fallacious. "[T]he faulty logic that equates infrequency with uniqueness" is not a logical fallacy.\textsuperscript{44}

The threshold issue for the law, therefore, is not the impossibility of falsifying universal propositions. It is whether criminalists are warranted in believing, as a practical matter, that certain universals (everyone has different fingerprints, everyone other than identical twins have different genomes, every face is unique, and so on) are true.\textsuperscript{45} If these beliefs are warranted, and if criminalists can measure the features that give rise to these differences with sufficient accuracy, then their claims to be able to individualize are sound. If these beliefs outstrip available theory and knowledge, as Saks and Koehler claim and as the NAS committee agrees (except, it

\textsuperscript{43} Id.

\textsuperscript{44} Id. at 205. A fallacy in logic is a faulty form of reasoning, not merely a conclusion that might be faulty. In a system of inductive logic, there is no formal fallacy in generalizing from repeated observations. Sample data permit inferences about population parameters—including the parameter that a characteristic occurs only once in the population. Naturally, there is a nonzero risk of error in accepting any inference about any population parameter. "The gap between the sample and the population will always require a leap of faith." PHILLIP I. GOOD & JAMES W. HARDIN, COMMON ERRORS IN STATISTICS (AND HOW TO AVOID THEM) 74 (2003). The only issue worth debating is the length of the leap.

\textsuperscript{45} Koehler and Saks read this sentence and personal correspondence not intended for publication as indicating that I use the words "establish the truth" to mean, in their words, "something . . . akin to a strong personal belief that has a solid foundation in data." Koehler & Saks, supra note 25, at 1196 n.35. In contrast, they define "establish the truth" to mean "that all point predictions other than the target prediction have been ruled out by the data," and they "simply do not accept [my] weakened definitional form of 'establish the truth.'" Id. This characterization prompts a brief clarification. Clearly, the warrant for a scientific belief or statement is not the strength of the conviction of the scientist who holds it. Neither does the warrant lie in vague allusions to "training and experience." Id. at 1197. Rather the question is whether the proposition has a secure foundation in data and theory. For most (and arguably all) forensic identification techniques, this foundation for strong claims of global uniqueness is missing. However, I do not embrace the rigid "definitional" view that a census is required to establish the likely truth (and what other kind of truth is there in science?) with respect to "all point predictions other than the target prediction." Id. at 1196 n.35. For example, if the probability of duplication really were as small as $10^{-41}$, see supra note 33, then the competing hypotheses effectively would have been ruled out, and the belief in the "target prediction" would be warranted.
seems, for DNA evidence),\textsuperscript{46} then the beliefs are either premature or false.

II. INDIVIDUALIZATION WITHOUT UNIQUENESS

Part I distinguished between claims of general uniqueness (no two pairs anywhere match) and specific uniqueness (no other object matches the particular trace seen in the case at bar). I argued that although general uniqueness is much more difficult to establish, an inductive proof of it is not beyond the capacity of science.\textsuperscript{47}

That said, the fact that general uniqueness is so hard to prove makes the traditional reasoning of many forensic-identification practitioners suspect, if not dogmatic. Rather than conduct the difficult empirical research that would be needed to establish that these objects or impressions are all uniquely identifiable, they postulate general uniqueness and use it to infer that a match then proves that a hair, a fiber, or the mark must have originated from the source that it matches. In the absence of proof of the premise of general uniqueness, however, this reasoning is insecure and might well be denominated an individualization fallacy.

Furthermore, the widely used distinction between “class” and “individual” characteristics\textsuperscript{48} encourages this individualization fallacy. The theory is that a large number of objects share class characteristics (such as shoe size), while other features (such as the scratches on the sole of a shoe) are individual characteristics. This definitional system promotes such tautologies as “[t]he uniqueness of an object may be established by an ensemble of individual [as opposed to class] characteristics.”\textsuperscript{49} Blithely postulating uniqueness in this manner, forensic science textbooks contain advice such as the following: “A positive identification is a conclusion that a

\textsuperscript{46} NRC 2009, supra note 7, at 87 (“[N]o forensic method other than nuclear DNA analysis has been rigorously shown to have the capacity to consistently and with a high degree of certainty support conclusions about ‘individualization’ . . . .”).

\textsuperscript{47} Additional analysis of conceptions of uniqueness and their relationship to individualization can be found in David H. Kaye, Identification, Individualization, and Uniqueness: What’s the Difference?, 8 L., PROBABILITY & RISK 85 (2009) (distinguishing between “universal individualization” and “local individualization,” between “general uniqueness” and “special uniqueness,” and between “universal uniqueness” and “local uniqueness”).

\textsuperscript{48} See, e.g., Thornton & Peterson, supra note 19, § 29:7, at 8 (describing the distinction and its use in the process of comparison).

\textsuperscript{49} Id.
particular shoe, and no other shoe, made the crime scene impression. No minimum number of individual identifying characteristics is needed to establish a positive identification.\textsuperscript{50} It would be better to eschew the class-individual distinction in favor of the realization that all characteristics are class ones. What matters is the size of the class. Some features are more discriminating than others. Shoe size, for example, is \textit{known} to be shared by many objects. It can be extremely valuable and easily used in excluding a given individual or object from further investigation.\textsuperscript{51} (If the shoe does not fit, you must acquit.) It also has rather modest probative value in showing that a specific individual or object is the source of the trace evidence. (Even if the shoe fits, you need not convict.) Other characteristics are far more variable. These supposedly “individual” characteristics pertain to a much smaller class (a class of one in the limit). As such, they have much greater probative value in establishing an association between two items than do the “class” characteristics that define a larger class.\textsuperscript{52}

But even if the traditional class-individual distinction begs the question of global uniqueness, a scientifically defensible opinion as to individualization is still attainable in some situations.\textsuperscript{53} This is the case for two reasons. First, testimony that a particular item is unique is a much weaker claim than testimony that all items are unique. Second, contrary to the loose or elliptical statements of many forensic scientists,\textsuperscript{54} individualization does not presuppose or imply even conditional uniqueness in a finite set. Suppose that a man suspected of stealing a jewel on a cruise ship has a fingerprint pattern that matches the latent print on the drawer from which the jewel was stolen and that there were 2500 people on board the ship at the time. Two thousand of them are fingerprinted, and only one is found to match. The other 500 people cannot be located. If the quality of the prints is high, the probability of a match to any of the 500 missing people could be

\textsuperscript{50} William J. Bodziak, \textit{Forensic Footwear Evidence, in FORENSIC SCIENCE: AN INTRODUCTION TO SCIENTIFIC AND INVESTIGATIVE TECHNIQUES} 297, 309 (Stuart H. James & Jon J. Nordby eds., 2003); \textit{see also id. at 298.}

\textsuperscript{51} \textit{See e.g., Thornton & Peterson, supra note 19, § 29:7, at 8.}

\textsuperscript{52} This perspective is congruent with Saks and Koehler’s emphasis on probability statements. \textit{See Saks & Koehler, supra note 13, at 211.}

\textsuperscript{53} For a more detailed discussion, see generally Kaye, \textit{supra note 47.}

\textsuperscript{54} \textit{E.g., Thornton & Peterson, supra note 19, § 29:10, at 11 (“Individualization implies uniqueness.”).}
miniscule. It then would be reasonable to conclude that the fingerprint examiner has identified the one individual who left the print. This is an individualization even though it does not imply that no one else on the earth has the same prints.\textsuperscript{55}

It might be thought that the cruise-ship example is contrived, but the point is more general. Following the invitation of the 1996 NRC report, FBI examiners, focusing on conditional uniqueness, have testified to source identifications in cases for which they consider the duplication probability for a particular profile in the United States population to be quite small.\textsuperscript{56} The 2009 report also seems comfortable with such testimony.\textsuperscript{57} Yet, the Saks-and-Koehler argument that any nonzero probability makes an assertion fallacious applies to DNA,\textsuperscript{58} and Fallacy brands an opinion that a defendant is the source of a DNA sample as “an evasion” of science.\textsuperscript{59}

\textsuperscript{55} Cf. Ian W. Evett & Bruce S. Weir, Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists 239 (1998) (“The issue for the forensic scientist is not ‘Is this profile unique’ . . . but ‘Is there sufficient evidence to demonstrate that they originate from the identical source.’”), Koehler and Saks object to using the word “individualization” to denote such testimony. They write that

The definition that Professor Kaye relies on reduces individualization to a subjective belief that is bolstered by evidence that falls far short of sufficient proof for this extreme claim. The difference between individualization as it is commonly understood and the definition offered by Professor Kaye is the difference between claiming that Alberto is the tallest man in the world because his measured height is greater than every other person in the world, and claiming that Alberto is the tallest man in the world either because an insufficiently tested theory assumes he is or because we have not seen anyone taller among those we have looked at.

Koehler & Saks, supra note 25, at 1201. This seems exactly backwards. Most critics maintain that it is the current definition of “individualization”—the “commonly understood” one described by Koehler and Saks—that produces testimony of identification based on subjective determinations that lack rigorous support. Relying on “subjective belief” and studies that fall “far short” is not a consequence of defining “an individualization” as a determination that one and only one individual is the source of the trace evidence in the case at bar. Similarly, this easily understood definition does not commit one to unsubstantiated claims about Alberto’s height or to “personal feelings or hunches.” Id. Neither Professors Koehler and Saks nor the last generation of forensic scientists can, by definitional fiat, confine the words “individualize” or “identify” to assertions based on the postulate of global uniqueness.

\textsuperscript{56} Roberto Suro, DNA Now Used to Make Specific Identification; FBI Calls Lab Match “Major Breakthrough”, WASH. POST, Nov. 13, 1997, at A4. Courts have upheld these source attributions despite defense arguments that untested relatives might match or that the laboratory could have erred. United States v. Davis, 602 F. Supp. 2d 658, 683-85 (D. Md. 2009); Young v. State, 879 A.2d 44, 56 (Md. 2005).

\textsuperscript{57} See supra note 46.

\textsuperscript{58} They also describe a different fallacy with regard to DNA evidence, exemplified by the prosecutor’s closing argument in People v. Simpson that if the random-match probability is one in 57 billion, the population must exceed 57 billion for a duplicate to exist. Saks & Koehler, supra note 13, at 203. As they cogently explain,
In sum, a well-founded and extremely tiny random-match probability indicates that, even if some other pairs of objects do match, the match at issue is not merely a coincidence; rather, it is a true association to a single source.\textsuperscript{59} In appropriate cases, therefore, it is ethical and scientifically sound for an expert witness to offer an opinion as to the source of the trace evidence. Of course, it would be more precise to present the random-match probability instead of the qualitative statement, but scientists speak of many propositions that are merely highly likely as if they have been proved. They are practicing rather than evading science when they round off in this fashion.

This is not to say that such testimony is the best method of communicating the test results to a lay jury. There is a cogent argument that such opinions are not helpful when more precise indications of probative value are available to permit the jury to reach its own conclusion about the source of the trace evidence.\textsuperscript{60} At this point, however, we are moving from what is scientifically acceptable to what is legally optimal. This

\textsuperscript{59} Saks & Koehler, \textit{supra} note 13, at 218 n.94. In doing so, they part company with the statistician whose exposition of “the uniqueness fallacy” motivates their putative fallacy. \textit{See} Balding, \textit{supra} note 34, at 148 (describing as “reasonable” the judgment that a characteristic is unique when it is “several orders of magnitude smaller” than the random match probability).

\textsuperscript{60} This inference assumes that other hypotheses (handling error, fraud, etc.) also have been eliminated. \textit{See} Kaye \textit{et al.}, \textit{supra} note 10, \textsection 12.3.1.

\textsuperscript{61} \textit{See}, e.g., Evett \& Weir, \textit{supra} note 55, at 241 (“Once we assign a [probability] then we must recognize that we have given the court something that they may choose to work with or without our assistance. Certainly, the idea that the scientist has some particular power to take that number and take a step equivalent to the Stoney ‘leap of faith’ is misconceived.”).
argument should be addressed on its merits and not foreclosed as an unscientific option.

III. PRESENTING LIMITED KNOWLEDGE IN COURT

The NAS committee recommends that “[f]orensic reports, and any courtroom testimony stemming from them, must include clear characterizations of the limitations of the analyses, including measures of uncertainty in reported results and associated estimated probabilities where possible.” 62 Saks and Koehler likewise implore criminalists to present “humbler, scientifically justifiable, and probabilistic conclusions.” 63 But neither group presents any serious analysis of whether numerical presentations are preferable to arguably more comprehensible qualitative ones, and what to do when probabilities are not at hand is open to debate. Inasmuch as there is a growing literature on these matters, I shall content myself with commenting on a few specifics, mentioning a variety of approaches, and pleading for eclecticism.

A. “Consistent-with” Testimony

Fallacy calls on criminalists who report that two samples are “a match” or “consistent” to add that this “does not require a conclusion that the patterns share a common source.” 64 To hammer home the point that other people or objects might match, Fallacy asks the witnesses to state that “in finding that two patterns match, they have placed the suspect object or person in a pool of one or more objects that match the evidentiary marks.” 65 The take-home message would be, “we have a match, but we cannot say how many other people or things would match.”

This “appropriate clarity and restraint” seems comparable to the weak presentation recommended by an NAS committee that reviewed the foundations for testimony about

---

62 NRC 2009, supra note 7, at 21-22.
63 Saks & Koehler, supra note 13, at 219.
64 Id. at 216; cf. Nat’l Research Council Comm. on Scientific Assessment of Bullet Lead Elemental Composition Comparison, Forensic Analysis: Weighing Bullet Lead Evidence 110 (2004) [hereinafter NRC 2004] (“The conclusions in laboratory reports should be expanded to include the limitations of compositional analysis of bullet lead evidence. In particular, a further explanatory comment should accompany the laboratory conclusions to readily portray the limitations of the evidence.”).
65 Saks & Koehler, supra note 13, at 216 (emphasis added).
the concentrations of various elements in bullet lead. The committee's findings brought to a halt this use of analytical chemistry to associate bullet fragments with boxes of ammunition. The committee had proposed limiting the formerly exuberant statements of analysts to testimony that bullets from the same compositionally indistinguishable volume of lead (CIVL) are more likely to be analytically indistinguishable than bullets from different CIVLs. An examiner may also testify that having... evidence that two bullets are analytically indistinguishable increases the probability that two bullets come from the same CIVL, versus no evidence of match status.

In colloquial terms, the committee recommended testimony to the effect that it is more likely to find a match when bullets come from the same blob of molten lead than when they come from different blobs. Such statements might reasonably be made with respect to other forms of trace evidence. A cautious analyst could report that similarities in head hairs are more likely when the hairs come from the same scalp than when they come from different scalps.

---

66 See NRC 2004, supra note 64.


68 In statistical jargon, this is a statement that the likelihood ratio or Bayes’ factor for the evidence exceeds one. See KAYE ET AL., supra note 10, §§ 12.4.2, 12.4.3. The 2009 NRC report implicitly endorses the use of more sharply defined likelihood ratios, at least for written reports. The report states that:

Although some disciplines have developed vocabulary and scales to be used in reporting results, they have not become standard practice. This imprecision in vocabulary stems in part from the paucity of research in forensic science and the corresponding limitations in interpreting the results of forensic analyses. Publications such as Evett et al., Aitken and Taroni, and Evett provide the essential building blocks for the proper assessment and communication of forensic findings.

NRC 2009, supra note 7, at 186 (citations omitted). The three authorities cited propose verbal scales for characterizing likelihood ratios. Two of them offer the following table for “reporting the value of the support of the evidence” (where V is the applicable likelihood ratio):

| $1 < V \leq 10$ | Limited evidence to support |
| $10 < V \leq 100$ | Moderate evidence to support |
| $100 < V \leq 1000$ | Moderately strong evidence to support |
| $1000 < V \leq 10000$ | Strong evidence to support |
| $10000 < V$ | Very strong evidence to support |
Neither of these approaches—the one-or-more-out-there statement nor the makes-it-more-likely assertion—clarifies how probative the evidence is. With bullet-lead comparisons, the jury might well be unduly impressed even with the modest statements of the results of inductively coupled plasma optical emission spectroscopy, which is the technique that gives the concentrations of the elements. This danger seems less when it comes to more mundane matters like the visual appearance of toolmarks and hairs. Thus, even though there is some risk of prejudice, this risk seems worth running with most forms of trace evidence, at least when compared with the alternative of entirely depriving the jury of a fair description of a relevant scientific finding.\footnote{70}

The proposals to scale back forensic science testimony reflect the view that an expert testifies not as an advocate, but as the representative of a learned profession, conveying its knowledge, along with its limitations, to the jury. A neutral expert not seeking to overawe the jury would express the important limitations up front. This is an attractive ideal but difficult to realize in practice. Attorneys often urge expert witnesses to suppress all qualifications and reservations—the joke about the desirability of the “one-handed economist,” the one who won’t say “on the other hand . . . ,” comes to mind.\footnote{71} Whether criminalists can stake out and preserve the independence to use both hands is far from obvious. Perhaps the NAS recommendation for an enforceable code of ethics would provide the requisite backbone. Although I am not sure that the law needs to require the expert to express every caveat in direct examination, much would be gained if the legal system or the forensic science profession insisted on written laboratory reports containing all the cautions. An analyst who holds these back on direct examination should be easy to impeach with his or her own report. In addition, a cautionary instruction from the judge might be of assistance.\footnote{72}

\footnote{70} See KAYE ET AL., supra note 10, § 12.5.1.

\footnote{71} Bartlett’s tentatively attributes the remark to President Truman. CLIFTON FADIMAN, & ANDRÉ BERNARD, BARTLETT’S BOOK OF ANECDOTES 542 (2000).

\footnote{72} See, e.g., NRC 1996, supra note 26, at 197.
B. Rarity and Numerical Testimony

Let us assume that the jury gets the message—a match is not an absolute identification. Can the criminalist do something more to explain its probative value? Obviously, this depends on what is known about the frequency of the identifying trait in the relevant population. Are the features very common, rarely seen, or somewhere in between? There will be occasions when such qualitative testimony is reasonable. When no duplicates have been seen after systematic, careful, and (one hopes) representative studies, a criminalist determined to refer to uniqueness might even assert that a trait is either unique or very rare in a population.

Numerical estimates should generally be possible, but not necessarily from the kind of research and modeling that Saks and Koehler describe. *Fallacy* suggests that forensic scientists can devise probability models for complex patterns of trace-evidence characteristics, sample the frequency of each characteristic in relevant populations, verify the independence of the characteristics, and multiply to arrive at random-match probabilities for consumption by juries. This will not be easy. Unlike nuclear DNA evidence, which both the 2009 report and *Fallacy* present as a model for all of forensic science, other patterns can be more complex, can vary substantially by locale, or can change over time.

Perhaps a more useful paradigm will turn out to be mitochondrial DNA sequences. These are used in cases where sample quantities are minute or the DNA is highly degraded. Mitochondrial DNA, which is found only in cytoplasm, reproduces asexually. Therefore, it is inherited from mother to child as a single unit. To give a numerical indication of the frequency of a lineage, the FBI maintains a collection of

---

73 See KAYE ET AL., supra note 10, § 12.5.2. A related proposal is to describe the strength of the evidence (the likelihood ratio) with phrases such as those listed supra note 69. Indeed, an expert witness with adequate information could give qualitative characterizations of both rarity and the likelihood ratio.

74 Saks & Koehler, supra note 13, at 217-18.


76 A small number of human hairs often contain enough mitochondrial DNA for sequencing to succeed. DAVID H. KAYE, THE DOUBLE HELIX AND THE LAW OF EVIDENCE 227-28 (2010).

77 The father merely contributes half of the nuclear DNA to the fertilized egg cell. *Id.* at 215. The cytoplasm is part of the mother’s egg cell. *Id.*
sequences from thousands of cases. It typically reports that the matching sequence in the case at bar has not been seen in this sample. For concreteness, suppose that the sample consists of 7000 sequences. When a defendant’s sequence is not in the database (which is the usual situation), one could say that so far it has been encountered one out of 7001 times. There is ample room to argue that this number should not be taken too seriously. It does not come from a random sample, and it might not be representative of the population in the vicinity of the crime. Still, it gives the jury some sense of how rare the sequence is, and jurors should be able to appreciate the limitations of the number, especially when there is discussion of how many other people in the vicinity of the crime might share the sequence.

With a mitochondrial-DNA-sequence database, each new sample is compared to all the previous ones to estimate the match frequency, and we know how many distinct sequences have been observed to date. Estimating the probability of a rare event from case work is trickier, but

---


79 See Pappas, 776 A.2d at 1104.

80 The upper confidence limit of the estimate also could be provided. The FBI does not do this. It quotes the upper bound of a 95% confidence interval above the proportion (0%) of the cases in which the sequence was previously seen. Id. at 1110 & n.11 (Conn. 2001). This is the “counting method” recommended in NRC 1992, supra note 22, at 75, to avoid any argument over the accuracy of multiplying individual probabilities with nuclear DNA.


82 Previous research by Koehler shows that when the expected number of matching individuals in a population is substantial, the “pool” formulation benefits defendants relative to the match-probability method. Jonathan J. Koehler, The Psychology of Numbers in the Courtroom: How to Make DNA-Match Statistics Seem Impressive or Insufficient, 74 S. CAL. L. REV. 1275, 1280-81 (2001). But the “pool” method also can lead to disputes over the size of the relevant population, see David H. Kaye, Rounding Up the Usual Suspects: A Legal and Logical Analysis of DNA Database Trawls, 87 N. CAR. L. REV. 425, 431 (2009) [hereinafter Trawls], and to fractional numbers of people, which could be difficult for jurors to understand. KAYE ET AL., supra note 10, § 12.4.1(2). When presented with both “pool” numbers and the match frequency, mock jurors were not overwhelmed by the mitochondrial DNA evidence. David H. Kaye et al., Statistics in the Jury Box: Do Jurors Understand Mitochondrial DNA Match Probabilities?, 4 J. EMPIRICAL LEGAL STUD. 797 (2007).
similar reasoning applies—and shows that failing to find any matches in a large number of comparisons of different objects can be quite informative.\textsuperscript{83} There are various methods for estimating the probability of an event that has occurred zero times in a sample of \( n \) observations—the “zero-numerator” problem.\textsuperscript{84} Applying the simplest of these, if a laboratory documented a zero numerator in 30,000 tests, it could infer a random-match probability of no more than 0.0001.\textsuperscript{85}

Naturally, there is a risk that the jury will draw too strong a conclusion from such a probability. This problem can be handled by cautioning the jury against inferring uniqueness or misconstruing the probability as a source probability. In addition, the quality of the estimate depends on the accuracy of the reported fraction of matching comparisons, the conditions under which the matches were made, and the analogy between a random sample and the casework or research sample.

\textbf{C. Source Testimony}

Experts in other fields routinely provide categorical statements. Pathologists opine as to the manner of death, psychologists to competence, and engineers to the cause of product failures. The law of evidence generally allows expert opinion testimony when it is well founded, but under normal relevance rules, existing theory and data on the discernible uniqueness of trace evidence typically are too weak to justify admission of an opinion that a pattern is unique.\textsuperscript{86} Contrary to

\textsuperscript{83} In assessing the probative value of a match, the sensitivity needs to be considered as well: How often do matches arise when the objects being compared are one and the same? See, e.g., V.L. Phillips et al., The Application of Signal Detection Theory to Decision-making in Forensic Science, 46 J. FORENSIC SCI. 294 (2001).


\textsuperscript{85} This estimate uses \( 3/n \) as an approximate upper bound on the probability. B. D. Jovanovic & P. S. Levy, A Look at the Rule of Three, AM. STATISTICIAN, May 1997, at 137. The “counting method” proposed for DNA typing, with a confidence coefficient of 0.95, see supra note 80, is equivalent to this simple rule.

\textsuperscript{86} An expert might still be justified in reporting the reasons for thinking that the match is probabilistic of identity. Statements such as “In my experience, different guns always give rise to different striations on the bullets” could be used in lieu of more precise studies, but unless the expert has meaningful experience in studying
recent federal district court opinions, experts should not be permitted to avoid the limitations in their knowledge simply by qualifying assertions of uniqueness with a fig leaf such as “to a reasonable degree of ballistic certainty.”

But what if the record were stronger? As long as there is a nonzero probability of duplication or another swan to consider—as there always will be—such testimony apparently fails Fallacy’s exacting standard for statistical proof. As we have seen, however, this austere standard does not comport with normal scientific practice. In the DNA field, scientists have indicated that opinions of general uniqueness or uniqueness of a particular DNA type within some smaller region are or will soon become scientifically acceptable.

To be sure, there is disagreement over how and whether to ascertain a precise value of a probability of uniqueness, and mismatches for guns that did not actually produce the striations, undocumented references to “experience” may reveal more about the analyst’s state of mind than they do about the state of the world. When testimony is of the “I know it, but neither I nor anyone else has studied it systematically” variety, a strong argument can be made that juries will overvalue the testimony and that exclusion is appropriate to encourage more extensive research.

87 United States v. Monteiro, 407 F. Supp. 2d 351, 375 (D. Mass. 2006) (holding that “the expert may testify that the cartridge cases were fired from a particular firearm to a reasonable degree of ballistic certainty” even though the accuracy of these judgments is unclear and “the expert may not testify that there is a match to an exact statistical certainty”); see also United States v. Mouzone, No. WDQ-08-086, 2009 WL 3617748, at *19-20 (D. Md. Oct. 29, 2009) (recommendations in magistrate’s report); United States v. Taylor, 663 F.Supp.2d 1170 (D.N.M. 2009); United States v. Diaz, No. CR 05-00167 WHA, 2007 WL 485967, at *14 (N.D. Cal. Feb. 12, 2007) (firearms toolmarks identification is admissible under Daubert, but “[t]he experts may not . . . testify to their conclusions ‘to the exclusion of all other firearms in the world.’ They may only testify that a particular bullet or cartridge case was fired from a particular firearm to a ‘reasonable degree of certainty in the ballistics field.’”).

88 See NRC 1996, supra note 26; NRC 1992, supra note 22; DANIEL L. HARTL & ANDREW G. CLARK, PRINCIPLES OF POPULATION GENETICS 131 (3d ed. 1997) (“Matches at 7 to 9 [VNTR] loci are virtually definitive of identity—barring technical errors in the DNA typing itself (such as mislabeling of blood samples) and except for identical twins.”); B.S. Weir, Discussion of “Inference in Forensic Identification”, 158 J. ROYAL STAT. SOC’Y A 49 (1995) (“[T]he chance that two unrelated individuals in a population share the same 16-allele [VNTR] profile is vanishingly small, and even for full sibs the chance is only 1 in very many thousands.”).


90 After reviewing this literature, one court agreed that “when a DNA method analyzes genetic markers at sufficient locations to arrive at an infinitesimal random match probability, expert opinion testimony of a match and of the source of the DNA evidence is admissible.” Young v. State, 879 A.2d 44, 45 (Md. 2005).

91 See David J. Balding, When Can a DNA Profile Be Regarded as Unique?, 39 SCI. & JUSTICE 257, 257 (1999) (“The probability that a defendant’s DNA profile is unique in a population of untyped individuals is . . . bounded below by one minus twice the sum of the match probabilities over the population . . . . However, because of the problem of the non-DNA evidence, there seems to be no satisfactory way for an expert
no canon of scientific reasoning demands that experts focus directly on the question of uniqueness.\textsuperscript{22} DNA experts can continue with the present regime of giving the incredibly small random-match probabilities that imply uniqueness. They can multiply these numbers by some arbitrary population size and report that in the “pool,” some number of people (even a fractional number) would be expected to have the defendant’s genotype. But the numbers could be a hair’s breadth away from the statement that no unrelated person would be expected to share this profile. Why not, as one astute statistician asks, give “a ‘plain English’ statement . . . ? For example, perhaps an expert witness could assert that, excluding identical twins and laboratory/handling errors, in his/her opinion the defendant’s

\textsuperscript{22} Cf. Saks & Koehler, supra note 13, at 218 n.94 (“[O]ffering source identifications at trial for sufficiently low probabilities would not be an implication of the science . . . .”). A scientific working group of the National Commission on the Future of DNA Evidence expressed the idea as follows:

The statistical basis for individualization is discussed by Evett and Weir (1998, pp. 243-244). The concept of individualization has been supported by Balding (1999). The FBI procedure has been criticized by Weir (1999) and supported by Budowle, Chakraborty, et al. (2000). Whether this, or in fact any statistical procedure for defining individualization is defensible continues to be debated. The procedure provides one way to interpret discriminatory power (a scientific question) in terms of “a reasonable degree of scientific certainty” (a subjective question). It is quite possible that within 5 years or less some such criterion will be accepted by the legal and forensic community, not as a scientifically appropriate statement, but as a practical definition for forensic purposes.

\textbf{NAT’L COMM’N ON THE FUTURE OF DNA EVIDENCE, RESEARCH & DEV. WORKING GROUP, THE FUTURE OF FORENSIC DNA TESTING: PREDICTIONS OF THE RESEARCH AND DEVELOPMENT WORKING GROUP 25-26 (2000).} The committee was explaining that there is no strictly scientific criterion for deciding how small the random-match probability must be to warrant an opinion that a particular genotype is “effectively unique.” It stated that:

Eventually, the probability becomes so small that the profile is effectively unique. The basis for concern would then be whether the techniques are adequate, the chain of custody is intact, the statistical treatment is appropriate, and no errors were made. But how small must such a probability be for a profile to be individualized?

\textit{Id.}
DNA profile is almost certainly unique in some appropriate population.\textsuperscript{93} Still other modes of presentation have been proposed.\textsuperscript{94} My objective is not to argue for any one of them, but only to ask that the issue be decided by examining the ease of presentation, comprehensibility, and scientific defensibility of all of them.

CONCLUSION

Radical skepticism of all possible assertions of uniqueness is not justified. Absolute certainty (in the sense of zero probability of a future contradicting observation) is unattainable in any science. But this fact does not make otherwise well-founded opinions unscientific or inadmissible. Furthermore, whether or not global uniqueness is demonstrable, there are circumstances in which an analyst can testify to scientific knowledge of the likely source of an object or impression. The admissibility of particular source attributions thus should turn on the actual state of this knowledge as applied to the task at hand, and to the helpfulness to the judge or jury of this testimony as contrasted with alternative presentations of forensically valuable findings. The optimal format for explaining the logical impact of a match is not self-evident. But it is clear that if forensic scientists are to contribute fully to the just resolution of criminal cases, they need a less absolutist and more nuanced theory of identification than the traditional presumption of characteristics that are intuitively judged to be individualizing.\textsuperscript{95} This is a fundamental—and fundamentally sound—message of the National Academy and of the broader academy.

\textsuperscript{93} \textbf{Balding, supra} note 34, at 136. In “the minority of cases in which uniqueness cannot reasonably be asserted,” Balding proposes “a probability of uniqueness.” \textit{Id. But see supra} note 91. A rough probability for uniqueness among unrelated individuals of the particular genotype in question could be provided in all cases. \textbf{See Kaye et al., supra} note 10, § 12.4.1(2).

\textsuperscript{94} \textit{See, e.g., Kaye et al., supra} note 10, § 12; \textit{Trawls, supra} note 81.

\textsuperscript{95} \textit{See, e.g., Christophe Champod \\ Ian W. Evett, A Probabilistic Approach to Fingerprint Evidence, 51 J. Forensic Identification 101 (2001); Saks \\ Kohler, supra} note 13.
Individualization Claims in Forensic Science: Still Unwarranted

Jonathan J. Koehler & Michael J. Saks

I. INTRODUCTION

In a 2008 paper published in the Vanderbilt Law Review entitled The Individualization Fallacy in Forensic Science Evidence, we argued that no scientific basis exists for the proposition that forensic scientists can “individualize” an unknown marking (such as a fingerprint, tire track, or handwriting sample) to a particular person or object to the exclusion of all others in the world.

In that article we made the following claims:

(1) the data necessary to achieve individualization have never been collected for any of the forensic science fields which aspire to individualize the source of crime scene evidence to its sole possible contributor;

(2) the best available—and perhaps the only scientifically defensible—approach to forensic identification is the use of random match probability estimates (which are not yet employed by any of the traditional forensic identification sciences);

(3) the argument that all objects are discernibly unique stands on little more than an oft-repeated maxim of forensic science legend and the illusory intuition that small frequencies imply uniqueness;

(4) probability estimates (by definition) cannot lead to uniqueness or individualization;

† Jonathan J. Koehler is Beatrice Kuhn Professor of Law, Northwestern University School of Law, 357 E. Chicago Avenue, Chicago IL 60611. jay.koehler@northwestern.edu. Michael J. Saks is Regents Professor, Sandra Day O’Connor College of Law and Department of Psychology, Arizona State University, Sandra Day O’Connor College of Law, 1100 S. McAllister Avenue, P.O. Box 877906, Tempe, AZ 85287-7906. michael.saks@asu.edu. The authors thank David Kaye for helpful comments.

(5) assertions of individualization generally exaggerate what is known or can be accomplished by forensic examiners.2

The central point and purpose of our article was a practical concern: to argue that because no field of forensic identification has adequate grounds for making individualization claims, expert witnesses from those fields should not make such claims in their reports and testimony. We recommended that, in the short term, expert witnesses should (a) revise their testimonial language to more accurately characterize the meaning and value of their findings, and (b) report only those inferences that can be supported by what is actually known by their fields.3 We further suggested that, in the long term, empirical research should be undertaken to place the forensic disciplines on more solid scientific footing.4

In this issue of the Brooklyn Law Review, we clarify, refine, and extend some of the ideas presented in Fallacy. Some of the refinements are prompted by Professor David Kaye’s paper, also in this issue of the Review,5 in which he takes issue with some of the arguments we made in Fallacy.

Moreover, we acknowledged that none of these essential insights is original to us. Others have discussed these problems for decades. We merely pulled these ideas together and discussed their implications.6 Others have also called attention to the difficulty or impossibility of justifying claims of individualization. See Fallacy, supra note 1, at 214-16. Some thoughtful forensic scientists, such as Christophe Champod and his colleagues, have responded by attempting to develop probabilistic characterizations of fingerprint comparisons. See Christophe Champod & Ian W. Evett, A Probabilistic Approach to Fingerprint Evidence, 51 J. FORENSIC IDENTIFICATION 101 (2001); Cedric Neumann et al., Computation of Likelihood Ratios in Fingerprint Identification for Configurations of Any Number of Minutiae, 52 J. FORENSIC SCI. 54 (2007). Others have begun to tame the language used when reporting the meaning of a match. See, e.g., Firearms and Toolmark Identification, in 4 MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY, § 34:1 (David L. Faigman et al. eds., 2008-2009).

Cf. NAT’L RESEARCH COUNCIL OF THE NAT’L ACADS., STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 7 (2009) [hereinafter NRC Report] (recognizing that none of the techniques that were the focus of our article have “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”); Lyn Haber & Ralph Norman Haber, Scientific Validation of Fingerprint Evidence Under Daubert, 7 LAW, PROBABILITY & RISK 87, 88 (2008) (arguing that the validity of the ACE-V fingerprint methodology has not been established through rigorous scientific experiments); Jennifer L. Mnookin, The Validity of Latent Fingerprint Identification: Confessions of a Fingerprinting Moderate, 7 LAW, PROBABILITY & RISK 127, 134 (2008) (“[T]he undeniable reality is that the community of forensic science professionals has not done nearly as much as it reasonably could have done to establish either the validity of its approach or the accuracy of its practitioners’ conclusions.”).

At the same time, we think it is important to point out that Professor Kaye appears to agree with our key points. For example, Professor Kaye does not believe that uniqueness has been established in any of the traditional, low-tech forensic sciences such as handwriting, toolmark identification, shoeprints, or fingerprints. He does not believe that mere matching, without a showing of uniqueness, can establish individualization in the typical case. He does not argue that testimony asserting that an object has been linked to its source to the exclusion of all others in the world is a scientifically reliable statement in any traditional forensic science discipline given the current state of knowledge. He does not dispute our claim that probabilistic statements rather than absolutist statements would provide a more accurate characterization of forensic identification. He agrees that there is a disconnect between the strong claims made by forensic scientists and the available scientific data. And he agrees that reform is in order.

If we agree on so much, what is the disagreement? The overarching difference is that while our focus leans toward the practical implications for courts of the large problems (on which we agree), Professor Kaye’s focus is on more abstract and conceptual issues—worthy of serious discussion, but with fewer implications for forensic science or judicial practice. Among

---

6 See generally id.
7 Although the heading of Section II in Professor Kaye’s article reads “Individualization Without Uniqueness,” it appears that he does not believe that individualization can be achieved in the typical case without uniqueness. Elsewhere he has written: “A true match establishes that the two samples of DNA have the same profile. Unless the profile is unique, however, a true match does not conclusively prove that the two samples came from the same source.” DNA Typing, in 4 MODERN SCIENTIFIC EVIDENCE, supra note 3, § 30:1 (emphasis added) (App. 30B, defining “True Match”).
8 “[U]nder normal relevance rules, existing theory and data on the discernible uniqueness of trace evidence typically are too weak to justify admission of an opinion that a pattern is unique.” Kaye, supra note 5, at 1182-83.
9 Some leading forensic scientists agree and have discussed this issue in detail. See Champod & Evett, supra note 3.
10 “[I]t is clear that if forensic scientists are to contribute fully to the just resolution of criminal cases, they need a less absolutist and more nuanced theory of identification than the traditional presumption of characteristics that are intuitively judged to be individualizing.” Kaye, supra note 5, at 1185-86.
11 Id. at 1165 (“With the imprimatur of the National Academy of Sciences behind recommendations for major change, the need for forensic scientists or analysts to retreat from the most extreme claims finally should be apparent to the judiciary as well as the forensic science community.”).
12 “I agree with the critique of a great deal of forensic science testimony . . . .” Id. at 1166.
these are: definitions (when can something properly be termed “metaphysical”?); logic and linguistics (when may probabilistic knowledge be expressed as an absolute? what are the customs of scientific communities for taking such linguistic shortcuts?); locating exceptions to general rules (are there current situations where an individualization claim is justifiable?); and thoughts about when an inferential leap might be small enough to be justified.  

13 Another possible difference of note is that many of Kaye’s illustrations and arguments are based on DNA or other normal sciences. However, Fallacy was concerned almost exclusively with non-DNA forensic identification: handwriting, tire impressions, shoe prints, fingerprints, toolmarks, firearms, etc. DNA databases and methods are not illustrative of how other forensic identification sciences operate. See NRC Report, supra note 4, at 7 (“With the exception of nuclear DNA analysis, however, 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.”).

Unlike most traditional forensic sciences, DNA typing stands out as an area that has, from its beginnings, respected the underlying probabilistic nature of forensic identification. Consider the following from a memo in which a crime laboratory informed the prosecutors with whom it works that it will cease using potentially misleading terminology and will limit its characterizations of the meaning of indistinguishably similar DNA samples to the associated random match probabilities:

The purpose of forensic DNA analysis is to determine whether an individual can be excluded as the donor of a body fluid stain or other bodily substance, the source of which is in question. Once a comparison has been made between the DNA profile results from the questioned source and the DNA profile results from the reference person, one of two conclusions may be drawn:

• the reference person is excluded as the donor of the questioned sample, or
• the reference person cannot be excluded as the donor of the questioned sample

This memo is to advise you that, effective November 1, 2003, the term ‘match’ will no longer be used in the conclusions of CFS DNA reports, in an effort to more clearly link the conclusion drawn from an analysis to its purpose. For example, when a DNA profile from person ‘A’ matches a DNA profile from the crime scene, the conclusion in the CFS report will state that “person ‘A’ cannot be excluded as the contributor of the crime scene profile.”

The significance of DNA findings will continue to be defined using a statistical calculation which addresses the probability of coincidentally selecting someone from the general population who also would not be excluded as the source of a DNA profile.

. . . .

The reported probability is the sole indicator of the significance of the finding that a person cannot be excluded as the source of a DNA profile.

Before responding to some of the issues that Professor Kaye raises, we offer a more specific definition of the “individualization fallacy” that we introduced in Fallacy. We also briefly compare and contrast this fallacy with others that appear in the literature.

II. DEFINING THE INDIVIDUALIZATION FALLACY

In Fallacy we did not offer a precise definition of the individualization fallacy. We try to remedy that here. The individualization fallacy refers to the belief that a particular known person or object must be the source of questioned markings whenever (a) the examiner judges that a sufficient number of characteristics are observable in both the questioned markings and the known, and (b) the examiner cannot otherwise distinguish the questioned markings from the known. In other words, the fallacy arises when the forensic scientist rules out all other possible sources for the unknown marking, including the multitude he has not examined, once he has found a single object or person that matches the features of the unknown marking. The fallacy is deeply entrenched in forensic science practice, where most examiners say that their knowledge, training, and experience enable them to make the inferential leap from observed consistencies between markings and their putative source to a conclusion that no other object in the world could have produced those markings.

Several subtleties and distinctions are worth noting concerning the notion of uniqueness and fallacies that are related to the individualization fallacy.

A. Uniqueness

A belief that one can individualize or has individualized is often bolstered by the claim that no two objects in the

14 Simon Cole, Against Uniqueness, Against Individualization, and how Wittgenstein Can Save Forensic Identification, LAW, PROBABILITY & RISK (forthcoming).

15 The International Association for Identification, which is one of the oldest and largest organizations of forensic science professionals, expects its members to offer fingerprint individualizations whenever matches are found. INT’L ASSOC. IDENTIFICATION, IAI POSITION CONCERNING LATENT FINGERPRINT IDENTIFICATION 1 (2007), http://www.theiai.org/current_affairs/fingerprint_position_paper_20071129.pdf (“The IAI endorses the position that individuals may be identified as the source of a particular friction skin impression . . . ”).
universe leave indistinguishably similar markings. This claim of uniqueness (first expressed in the maxim that “nature never repeats”) is often proffered by forensic science practitioners, forensic science authorities, courts, and even federal agencies as “a defensible epistemological foundation for forensic testimonial claims of source attribution.”

Assuming uniqueness to be true shortens the inferential chain from the perception that two markings are indistinguishably alike to the claim that whatever made one set of markings must have made the other.

But uniqueness is not essential to the practice of forensic individualization. Some examiners assert that they need deal only with the samples in front of them—the questioned and the known—and need give no thought at all to the frequency with which particular characteristics or sets of characteristics exist within the population from which those mark-producing objects came. Others claim to be able to discern “individualizing” characteristics from non-individualizing markings, again without concerning themselves with population distributions of mark-producing objects. To be sure, there are certain circumstances where individualization could be achieved without having to make reference to the full population of such objects. But such circumstances are exceptional and do not explain why practitioners in virtually all areas of forensic identification—with the notable exception of DNA typing—behave as if the population of potential sources is of no consequence to the task of individualizing.

---

16 Cole, supra note 14, at 12.
17 Compare the reasoning of examiners of fingerprints, firearms, handwriting, etc., to that of DNA examiners. The former go directly from the perception of great similarity to a conclusion of individualization. The latter must pause to calculate random match probabilities based on population data.
18 Cole, supra note 14, at 26 (“What distinguishes areas of friction ridge skin from these other objects is not ‘uniqueness’; it is their diagnosticity: our ability to assign traces of these objects to their correct source with a certain degree of specificity under certain parameters of detection and under certain rules governing such assignments.”).
19 See Kaye, supra note 5, at 1173-77.
20 This disregard of populations might be the result of the underlying assumption of uniqueness being so fully incorporated into forensic individualization practice that examiners have forgotten, or set aside, the argued basis for ignoring populations.
B. The Uniqueness Fallacy

In Fallacy, we mentioned that the individualization fallacy is a cousin of David Balding’s uniqueness fallacy. The uniqueness fallacy is the mistaken belief that whenever the expected number of people or objects sharing a set of known characteristics is less than one, then one may infer that the known person or object is unique. Some proponents of individualization have made, and many courts have accepted, the argument that if the population is smaller than the inverse of the random match probability, then uniqueness is established. For example, if there are 6 billion people on earth, and an analyst reports that the relevant random match probability for, say, a DNA profile is 1 in 20 billion, then it is fallacious to conclude that the DNA profile in question must be unique. Fallacy cites Balding to explain why this is a fallacy, and presents other illustrative explanations.

C. The Fingerprint Examiner’s Fallacy

According to Simon Cole, the fingerprint examiner’s fallacy occurs when the (assumed) uniqueness of fingerprints is invoked to support the asserted accuracy with which fingerprint examiners can identify the source of latent prints. As Cole points out, the relationship between fingerprint uniqueness and examiner accuracy is a tenuous one. Even if fingerprints are unique, it is fallacious to assume that uniqueness somehow confirms the accuracy of examiners’ identifications. By analogy, the (assumed) fact that every

21 Fallacy, supra note 1, at 205; see David J. Balding, Weight-Of-Evidence: DNA Profiles 148 (2005).
22 Fallacy, supra note 1, at 203.
23 Id. at 203-05.
24 Simon A. Cole, Grandfathering Evidence: Fingerprint Admissibility Rulings from Jennings to Llera Plaza and Back Again, 41 AM. CRIM. L. REV. 1189, 1198 (2004) (defining this fallacy: “The fingerprint examiner’s fallacy consists of reasoning that the uniqueness of the object of forensic study vouches for the validity of a forensic matching process.”). In the course of explicating the “fingerprint examiner’s fallacy,” Cole summarizes other elements that we develop in Fallacy and which others have raised for a long time. Indeed, Cole’s historical work shows that the problems associated with making justifiable individualization claims have been appreciated by forensic scientists and scholars for at least a century. Id. at 1199 (citing Henry Faulds, Guide to Finger-Print Identification 51 (1905)). However, this awareness has largely been hidden from the courts, which may help explain why the NRC Report, supra note 4, at 53, concludes that “the courts have been utterly ineffective” in filtering forensic identification evidence.
25 Cole, supra note 24, at 1198.
human face is distinguishable from every other human face does not assure that eyewitness identifications are always accurate. Cole argues that much more is involved in drawing conclusions and in evaluating the risk of error, and that claims of uniqueness do not get us very far in those regards.  

In sum, several related forensic science fallacies have been identified, but they are distinct. The uniqueness fallacy concerns the faulty reasoning that match probabilities smaller than the reciprocal of the population of interest lead to inferences of uniqueness. The fingerprint examiner’s fallacy concerns the faulty reasoning that turns the alleged uniqueness of fingerprints into an argument for the accuracy of fingerprint identifications. The individualization fallacy, which was suggested in Fallacy, concerns the faulty reasoning that certain observations are sufficient to individualize, regardless of whether uniqueness is invoked in support.

III. RESPONDING TO PROFESSOR KAYE

Professor Kaye takes issue with what he refers to as our “radical skepticism of uniqueness,” though, as mentioned in the Introduction, he agrees with the gravamen of Fallacy. The disagreements that Professor Kaye has with Fallacy have more to do with theoretical issues about what might be possible in forensic science than with practical concerns about what has been achieved and how those achievements are reflected in courtroom practice. Three topics on which we do appear to disagree are metaphysics, uniqueness, and individualization.

A. Metaphysics

Professor Kaye questions our suggestion that forensic individualization rests more on metaphysical and rhetorical grounds than on scientific and empirical grounds. As we understand his argument, Professor Kaye’s position is that the individualization hypothesis is not metaphysical because “individuality . . . [concerns] measurable characteristics that

26 Id. at 1201-03.
27 We doubt that our position on uniqueness qualifies as “radical skepticism.” Our position is that it has not been proved and that it seems unlikely that so extreme a position as individualization-to-the-exclusion-of-all-others—a frequently-invoked foundation stone of forensic individualization—could be proved. Kaye, supra note 5, at 1166.
28 “Although I agree with the critique of a great deal of forensic science testimony . . . .” Id.
can exhibit unequivocal differences and similarities.” Although we largely agree with the quoted statement, it reflects only part of the picture. The larger and more pragmatic part is that individualization claims come to court supported mainly by exaggerated rhetoric and reasoning that is grounded in little empirical data. Professor Kaye agrees that the rhetoric of forensic individualization far exceeds the science, though he takes issue with our use of the word “metaphysical” to describe the foundations for the individualization claim. We maintain that our use of the term in this context is accurate and appropriate.

In a concluding section in Fallacy titled “Unproved and Perhaps Unprovable,” we wrote:

In sum, no sound and rigorous evidence supports the assumption of unique individualization. Moreover, the assumption is so heroic and the research required to test it seriously would be so massive that one must doubt whether it is possible to conduct an empirical study or set of studies that would provide solid support for the hypothesis.

Our claim, then, is not that there is a rule of logic or ontology that prevents individualization. Instead, we are concerned with the more practical issues of whether individualization has been proven, how amenable it is to testing, and whether the self-presentation of these fields in court accurately reflects the limitations of testing, proof, and case-specific conclusions. On all of these issues, we think forensic individualization science has fallen short despite being in the expert witness business for a century. Consequently, we think it is fair to characterize the individualization claim as predominantly rhetorical. As for our invocation of

29 Id. at 1167.
30 “Unequivocal” is too strong. Perceptual and judgmental disagreements will be common in the process of making such assessments. See, e.g., I.W. Evett & R.L. Williams, A Review of the Sixteen Points Fingerprint Standard in England and Wales, 46 J. FORENSIC IDENTIFICATION 49 (1996).
31 NRC Report, supra note 4, at 188-89 (“In most forensic science disciplines, no studies have been conducted of large populations to establish the uniqueness of marks or features”); id. at 184 (“[T]he concept of ‘uniquely associated with’ must be replaced with a probabilistic association . . . .”); see also Harry T. Edwards, Solving the Problems That Plague the Forensic Science Community, 50 JURIMETRICS 5, 8-9 (2009) (referring to the “dearth of scientific research to establish limits of performance, to ascertain quantifiable measures of uncertainty, and to address the impact of the sources of variability and potential bias in fingerprint examinations and in other forensic disciplines that rely on subjective assessments of matching characteristics”).
32 See supra notes 8, 10-12 and accompanying text.
33 Kaye, supra note 5, at 1167.
34 Fallacy, supra note 1, at 213-14.
“metaphysical,” readers who consult a good dictionary are likely to find that there are several meanings of this word that fit well with the situation we have described.

B. Uniqueness

Professor Kaye seems to agree with our central point about uniqueness. That is, he seems to agree that data—either in principle or in practice—cannot establish the truth of a uniqueness point prediction. He says: “Uniqueness means that the proportion of objects with the given feature in the whole population of size \( N \) is exactly \( 1/N \). Yet, no matter how close the sample proportion comes to \( 1/N \), the next sample datum could establish that the population proportion is \( 2/N \) or more.”

However, Professor Kaye does not think that the fact that empirical sampling has not and cannot demonstrate uniqueness is the important point.

Email from David Kaye to Jay Koehler & Michael Saks (Feb. 21, 2010, 15:16) (on file with authors) (emphasis added). In this same correspondence, Professor Kaye also says that he believes that “one can ‘establish the truth’ of a proposition without being 100% certain that it is true.” Id. Apparently, then, our disagreement turns on what it means to “establish the truth” of a proposition. When we say “establish the truth” of a particular point prediction, we mean that all point predictions other than the target prediction have been ruled out by the data. We interpret Professor Kaye’s use of the phrase “establish the truth” to mean something substantially weaker, something more akin to a strong personal belief that has a solid foundation in data. Although language is often sufficiently imprecise that there is room for personal interpretation, we simply do not accept this weakened definitional form of “establish the truth.” Nor do we believe that this definition fits with a common understanding of the phrase. For example, we suspect that if an examiner claims that data have “established” that such-and-such is true, jurors will not interpret this to mean merely that the examiner has a strong belief in the proposition and has some data to back up this belief. Instead, jurors are likely to equate the examiner’s establishment claim with the indefensible claim that the data have established the claim to a certainty. In other words, there is no practical difference between a claim that data have established the truth of a uniqueness point prediction, and a claim that data have established that truth to an absolute certainty. And this is why we say that, linguistic preferences aside, there is no meaningful difference between our position and Professor Kaye’s on the issue of whether data can truly establish uniqueness: we all agree that they cannot.

---

**Note:**

35 Professor Kaye objects to this characterization of his position. That is, whereas we say that he agrees with us that data cannot establish the truth of a uniqueness point prediction, he says that his position is that “data cannot establish to an absolute certainty the truth of any point estimate of any population parameter.” Email from David Kaye to Jay Koehler & Michael Saks (Feb. 21, 2010, 15:16) (on file with authors) (emphasis added). In this same correspondence, Professor Kaye also says that he believes that “one can ‘establish the truth’ of a proposition without being 100% certain that it is true.” Id. Apparently, then, our disagreement turns on what it means to “establish the truth” of a proposition. When we say “establish the truth” of a particular point prediction, we mean that all point predictions other than the target prediction have been ruled out by the data. We interpret Professor Kaye’s use of the phrase “establish the truth” to mean something substantially weaker, something more akin to a strong personal belief that has a solid foundation in data. Although language is often sufficiently imprecise that there is room for personal interpretation, we simply do not accept this weakened definitional form of “establish the truth.” Nor do we believe that this definition fits with a common understanding of the phrase. For example, we suspect that if an examiner claims that data have “established” that such-and-such is true, jurors will not interpret this to mean merely that the examiner has a strong belief in the proposition and has some data to back up this belief. Instead, jurors are likely to equate the examiner’s establishment claim with the indefensible claim that the data have established the claim to a certainty. In other words, there is no practical difference between a claim that data have established the truth of a uniqueness point prediction, and a claim that data have established that truth to an absolute certainty. And this is why we say that, linguistic preferences aside, there is no meaningful difference between our position and Professor Kaye’s on the issue of whether data can truly establish uniqueness: we all agree that they cannot.

36 Kaye, supra note 5, at 1170-71.

37 Professor Kaye objects to this characterization of his position for reasons similar to those described supra note 35. He thinks that sampling “can demonstrate uniqueness (in principle, in some populations) especially when considered together with an understanding of the sources of randomness.” Email from David Kaye to Jay Koehler & Michael Saks (Feb. 21, 2010, 15:16) (on file with authors) (emphasis added). Once again, we confront a linguistic problem where we and Professor Kaye are not using common words and phrases (like “demonstrate uniqueness”) in the same way.
the focus should be on “whether criminalists are warranted in believing” that fingerprints are unique.\textsuperscript{38}

We agree that the uniqueness question must turn on the issue of what the science supports, but it is not clear that we and Professor Kaye draw the same inferences from that science. In our view, the existing and foreseeable scientific knowledge falls far short of providing criminalists with enough scientific support to claim that the objects that they study are either unique or discernibly unique. Certainly the uniqueness question cannot turn on the beliefs that forensic scientists have about this issue based on their training and experience.

Throughout most of the twentieth century, criminalists may have felt that they had good justification for believing many things that subsequent scientific study revealed to be untrue. For example, at one time criminalists believed that they were impervious to context effects, or that bullets with similar trace element profiles necessarily had been manufactured in the same lead melt. But subsequent scientific study indicated that those beliefs were either untrue\textsuperscript{39} or premature.\textsuperscript{40} People hold sincere but mistaken and unsupported beliefs all the time and, in many cases, the negative consequences are minimal. But when representatives of an assertedly scientific discipline allow assumption and good faith belief to substitute for good grounds\textsuperscript{41} and scientific knowledge,\textsuperscript{42} the practical effects can hamper the advancement of the science as well as the search for justice in particular cases. When forensic scientists testify under oath that markings are unique, they rarely qualify that testimony by conceding that the claim reflects more of a personal belief than a scientific fact.\textsuperscript{43} The practical reality is that most forensic scientists

\textsuperscript{38} Kaye, supra note 5, at 1172.

\textsuperscript{39} NAT'L RESEARCH COUNCIL, FORENSIC ANALYSIS: WEIGHING BULLET LEAD EVIDENCE 96-99 (2005); Itiel E. Dror et al., Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications, 156 FORENSIC SCI. INT'l 74, 77 (2006) (“[I]t is possible to alter identification decisions on the same fingerprint, solely by presenting it in a different context.”).

\textsuperscript{40} NAT'L RESEARCH COUNCIL, BALLISTIC IMAGING 3 (2008) (Daniel L. Cork et al. eds. 2008) (“The validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated.”).

\textsuperscript{41} “Proposed testimony must be supported by appropriate validation—i.e., ‘good grounds’ . . . .” Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 590 (1993).

\textsuperscript{42} “[S]cientific’ implies a grounding in the methods and procedures of science. Similarly, the word ‘knowledge’ connotes more than subjective belief or unsupported speculation.” Id.

\textsuperscript{43} See supra note 31 and accompanying text.
bolster the perceived probative value of their individualization testimony at trial by asserting that evidentiary markings are unique, and that forensic scientists can individualize by discerning that uniqueness. This is bad science, bad policy, and should not be welcome on grounds that the testifying expert is merely expressing a sincere belief.

C. Individualization

Professor Kaye takes exception to our skepticism about whether individualization claims can be proved. In Fallacy, we argued that individualization—the process of linking an unknown marking to a source, to the exclusion of all other possible sources—is “unproved and perhaps unprovable.” Professor Kaye disagrees. But much of significance can be said about our disagreement, starting with what he means by the term “individualization,” which departs from the conventional meaning in important ways.

1. Definitions

As we and forensic scientists themselves use the term, individualization refers to a finding that a particular print or marking was produced by a particular source, to the exclusion of all other possible sources on the planet. Importantly, individualization is not merely a conclusion that a particular source might be the source of a target marking, that many other possible sources can be ruled out as the source, or that there is only a slim chance that any source other than the named one would share the observed characteristics of the unknown marking. Individualization is a claim that all potential sources but one have been affirmatively ruled out as the person or object that produced the print or marking in

---

44 Scientific Working Group on Friction Ridge Analysis (“SWGFAST”), 2009 Standards for Minimum Qualifications and Training to Competency for Friction Ridge Examiner Trainees (fingerprint examiner training document stating that examiners must “[u]nderstand the basic foundations for friction ridge examination (persistence and uniqueness) as a means of identifying the source of an impression”); see also NRC Report, supra note 4, at 43 (forensic scientists “believe that unique markings are acquired by a source item in random fashion and that such uniqueness is faithfully transmitted from the source item to the evidence item being examined”).


46 Fallacy, supra note 1, at 208.
question. That is the definition of individualization most widely understood within the forensic sciences.\footnote{The individualisation of an impression is established by finding agreement of corresponding individual characteristics of such number and significance as to preclude the possibility (or probability) of their having occurred by mere coincidence, and establishing that there are no differences that cannot be accounted for.} Forensic Human Identification: An Introduction 74 (Tim Thompson & Sue M. Black eds., 2007) (quoting Harold Tuthill, Individualization: Principles and Procedures in Criminalistics (1994)); see also 4 Modern Scientific Evidence, supra note 3, § 30:19; NRC Report, supra note 4, at 43-44; Keith Inman & Norah Rudin, The Origin of Evidence, in 126 Forensic Sci. Int'l 11, 11-16 (2002).

In the conventional practice of fingerprint examiners, a different term, “identification,” is considered a proper way to express a conclusion of “individualization.”\footnote{See SWGFAST, Quality Assurance Guidelines for Latent Print Examiners (ver. 3.0, 2006) available at http://www.swgfast.org/Quality_Assurance_Guidelines_for_Latent_Print_Examiners_3.0.pdf (referring to “individualization” throughout); Christophe Champod, The Inference of Identity of Source: Theory and Practice, Address at the First International Conference on Forensic Human Identification in The Millennium, at 1 (Oct. 1999), available at http://www.latent-prints.com/images/The%20Inference%20and%20Identity%20of%20Source.pdf (“Among identification fields, the term identification generally denotes individualisation”). However, identification is also used by forensic scientists to refer to the process of determining the category to which an object belongs. This is the nature of the answer to the question, “what chemical substance is this white powder?” See, e.g., 4 Modern Scientific Evidence, supra note 3, § 30:19; NRC Report, supra note 4, at 36; Inman & Rudin, supra note 47.} As Thornton and Peterson explain:

[In everyday usage, the term identification often is used when the concept of individualization is intended. One may hear testimony of the sort, “I identified the latent fingerprint as having been made from the right ring finger of the defendant.” The intent of the witness here is to declare clearly that the latent fingerprint was that of the defendant, to the total exclusion of all other fingers of all of the other people in the world. The use of the term “identified” here is not the most precise usage of the word; the term “individualized” would be more felicitious. But use of the term “individualization” and various other forms of the word would only confuse matters. If, in response to the question, “Did you have occasion to identify the suspect’s fingerprint on the knife?” the witness were to answer, “No, I individualized it,” communication would be thwarted and the listener confused.\footnote{4 Modern Scientific Evidence, supra note 3, § 30:19.}]

Although most examiners use the terms individualization and identification interchangeably—a practice that is sometimes promoted by forensic science working groups\footnote{See SWGFAST, supra note 3.}—some fingerprint examiners have recently suggested a novel distinction between individualization and
identification. This latest redefinition is significant for our purposes here because it may help explain how Professor Kaye arrives at his own definition of individualization.

In a recent Minnesota case, two forensic scientists testified that individualization claims cannot be proven: “The only way you could prove that is to look at every single person on the planet. It’s not physically possible and even if you could do that, that doesn’t mean someone tomorrow won’t be born with the same friction ridge skin.” One of these examiners went on to draw a distinction between an individualization and an identification. According to this examiner, “identification” means that the examiner

made a decision that the chance that someone else could have left [the print] is so remotely small, he’s willing to dismiss it and say yes, I believe that this latent print in my opinion was produced by that individual. He did not say that he’s excluded everyone else on the planet and he left a theoretical possibility that there might be someone else on the planet that could have produced a similar looking latent print. And he has no way of calculating what that probably is at this time.

One explanation for this novel attempt to distinguish between an individualization and an identification might be that it is an attempt to have one’s cake and eat it too. Examiners can, on the one hand, acknowledge the scientific impossibility of “individualizing” while at the same time preserve the ability to pinpoint a suspected source by asserting an “identification” and standing it on a more humble foundation of personal opinion.

Professor Kaye defines individualization much as examiner Langenburg seeks to define the weaker term identification: as “the conclusion that ‘this trace came from this

---

51 Transcript of Record at 148, State v. Hull, 2008 WL 4301902 (Minn. Ct. App. Aug. 28, 2008) (CR-07-2336) (cross-examination of Glenn Langenburg). A second forensic scientist testifying in this same hearing testified that, “the only way to really say an individualization could occur, is to actually do comparisons to all prints of everyone that has ever lived.” Id. at 48-49 (direct examination of Joshua Bergeron); see also Christophe Champod, Identification and Individualization 1 (Nov. 6, 2008) (unpublished manuscript) (concluding that “individualization conclusions are out of reach and cannot be easily substantiated, either in the classic identification fields (such as fingerprint evidence) or in DNA profiling”).

52 Transcript of Record at 149, Hull, 2008 WL 4301902 (CR-07-2336).

53 Champod & Evett, supra note 3, at 103 (demonstrating that “the inferential process of identification . . . is essentially inductive and hence probabilistic”).
individual or this object.” But individualizations are more than subjective source conclusions or a witness’s personal feelings or hunches. They are bold statements about the world that require proof that cannot now (and probably never will be) obtained. The definition that Professor Kaye relies on reduces individualization to a subjective belief that is bolstered by evidence that falls far short of sufficient proof for this extreme claim. The difference between individualization as it is commonly understood and the definition offered by Professor Kaye is the difference between claiming that Alberto is the tallest man in the world because his measured height is greater than every other person in the world, and claiming that Alberto is the tallest man in the world either because an insufficiently tested theory assumes he is or because we have not seen anyone taller among those we have looked at.

2. Small Population Examples

Professor Kaye posits that “there are circumstances in which an analyst reasonably can testify to having determined the source of an object.” He offers two such circumstances. In one, he describes an unusual situation where a fingerprint examiner likely would be justified in claiming that he individualized a latent print. The situation concerned a crime that was known with certainty to have occurred at sea and a latent print that was known with certainty to belong to one of a relatively small number of passengers, all of whom are available for testing. We agree that there are some circumstances in which the potential source population of a print or marking may be narrowed to a small, accessible set. Indeed, such examples are not unfamiliar in discussions of individualization. However, these examples provide more

54 Kaye, supra note 5, at 1166.
55 Id.
57 One of the authors has used it himself. See Michael J. Saks, Explaining the Tension Between the Supreme Court’s Embrace of Validity as the Touchstone of Admissibility of Expert Testimony and Lower Courts’ Rejection of Same, 5 EPISTEME 329, 342 n.1 (2008) (“There is at least one circumstance where certainty of individualization could be achieved: If the candidates for the perpetration of a crime could be narrowed to a finite group, and each member of the group had distinguishable markings (be they fingerprints or something else), and it was known that the
support for the argument that individualization claims are generally unwarranted than they do for the argument that individualization should be broadly encouraged. Small, closed population examples “work” only because one can compare target latent prints to every member of the potential source population. The presence of this unusual circumstance is what sets the stage for an individualization claim (provided, of course, that all but one print can be eliminated as potential sources of the target latent).

But what about the more typical situation where the potential source population cannot be narrowed much beyond the general population or some other large population? In these cases, examiners are not able to eliminate every member of the potential source population, and therefore they are not about to identify a source using the logical rigor that arises in the small, closed population example. With this in mind, the relevance of the ship hypothetical that Professor Kaye offered is not to say “if individualization can be achieved in this context it can be achieved in others.” Instead, these types of hypotheticals remind us that the defensible approach of comparing a questioned print or marking to all prints or markings in the potential source population is often not possible. And when forensic examiners can do no better than sample from larger potential source populations and draw inferences from their findings, then they must forsake absolutes. In its place, forensic scientists should do what other scientists do: offer suitably cautious conclusions that make use of the tools of probability and statistics.  

3. Small Random Match Probabilities and the Inferential Leap

Professor Kaye argues that a second circumstance in which claims of unequivocal source and individualization claims are justified occurs when random match probabilities

\[ \text{perpetrator left his markings on the body (or on whatever), then the person who matched the crime scene markings would have to be the perpetrator.)}. \]

\[ \text{See Champod & Evett, supra note 3. We should also note that a new breed of sophisticated forensic scientists working in the fingerprint area is developing models and procedures aimed at providing transparent, empirically based, probabilistic conclusions to replace individualization claims. Glenn Langenburg & Cedric Neumann, Moving Towards Using Statistics for Fingerprint Evidence in the Courtroom, Proceedings of the American Academy of Forensic Sciences, Seattle, WA, February 26, 2010.} \]
are very small." In a nutshell, his argument is that object uniqueness and individualization are, for all intents and purposes, proved even when there remains a chance that some objects are not unique or that an individualized marking was actually produced by some unexamined object. Science, he says, does not require absolute certainty. Instead, we draw reasonable inferences from the data we have and proceed as if that inference were absolutely true.

Professor Kaye offers, as an example, the treatment of Ohm's law: "Ohm's law might not be exactly right, or it might break down tomorrow, but electrical engineers can safely assume that it is absolutely true." The implication is that even if forensic examiners can't be 100% sure of their ability to individualize, they are safe in proceeding on the assumption that their individualization conclusions are absolutely true.

An easy response is that Ohm's law not only "might not be exactly right," but it actually is demonstrably wrong under so many conditions that electrical engineers are not safe to assume that it is "absolutely true." The implication of the conditional nature of Ohm's law for forensic science individualization claims is that it would not be safe for courts to simply regard individualization as absolutely true. But this is too easy. The example Professor Kaye chose was flawed, but we trust that there exist better illustrations of the point he wishes to make.

Our response to those better examples would be that even the best of them is not an apt analogy to the problem that forensic individualization presents to courts. First, engineers

59 Kaye, supra note 5, at 1176 ("[A] well founded and extremely tiny random-match probability indicates that, even if some other pairs of objects do match, the match at issue is not merely a coincidence; rather, it is a true association to a single source. In appropriate cases, therefore, it is ethical and scientifically sound for an expert witness to offer an opinion as to the source of the trace evidence." (footnote omitted)). Arguably, that final inferential leap is for the factfinder to make—using knowledge supplied by the expert—not for the expert to make for the factfinder, with no basis greater than what the factfinder now has. See Wells, infra note 77 and accompanying text.

60 Kaye, supra note 5, at 1174.
61 Kaye, supra note 5, at 1168.
62 Ohm's law—which states (in part) that the current between two points in a conductor is directly proportional to the potential difference (voltage) across the two points—"holds only approximately and under limited conditions and not for all materials." SCIENCE AND TECHNOLOGY ENCYCLOPEDIA 376 (1999). Although there are contexts within which electrical engineers may safely treat that the law as true, there are other contexts where such an assumption would spell disaster. Forensic individualization assumptions probably operate the same way, except that we know far less about the conditions under which they do and do not hold.
and scientists have more intimate knowledge of their theories and data, and appreciate their limitations. Lawyers, judges, and jurors are much less likely to understand the limitations of the claims being made for forensic individualization, and the limitations of its theory, its data, and its conclusions. Second, the central claim of forensic individualization science is qualitatively different from all or virtually all other sciences. Where other sciences cautiously test hypotheses about relationships among variables, the forensic individualization sciences simply offer error-bar-free point predictions without backing those predictions with anything that approaches scientific validation. At the risk of redundancy, it is important to be clear about what individualization is and is not. It is not simply a scientific classification claim in which the object in question is one of a small number of possible sources for the questioned marking. Nor is it a claim that the probability that another object is the source is low or even extremely low. It is a claim that the probability that an object other than the one identified by the examiner could be the source of the questioned marking is exactly zero. The examiner justifies his impossibility thesis by making an unwarranted inferential leap from the mere observation of similar markings or, at best, an impression or intuition of a low frequency of such markings.

63 Summarizing a central theme in the NRC Report, one of the co-chairmen of the report refers to the “paucity of scientific research to confirm the validity and reliability of forensic disciplines.” See Edwards, supra note 31, at 2.

64 The notion that forensic individualization claims are extreme and fundamentally unscientific is neither a radical idea nor one that is original with us. Consider, for example, the following passage from a 1998 book by the highly respected forensic statistician Ian Evett and equally respected biostatistician Bruce Weir about the meaning of a fingerprint “identification”:

We should be in no doubt about the degree of certainty implicit in a fingerprint identification. The expert is, in effect, saying “I am certain that this latent mark and this control print were made by the same person and no amount of contrary evidence will shake my certainty”. Or, to look at this from a Bayesian perspective, no matter how small the prior odds are, the likelihood ratio is so large that the posterior odds approach infinity. Stoney sees that a fingerprint identification is based on a “leap of faith,” and he is quite correct to conclude that such a leap of faith has nothing to do with scientific principles. It is that leap of faith that characterizes the essence of a conclusion of identity of source and, as he points out, that is a fundamental difference between fingerprint evidence and DNA evidence. Stoney’s “leap of faith” is equivalent to attaining an infinite likelihood ratio; this kind of belief cannot derive from any scientific process.

One must remain mindful of where these shortcuts—these inferential leaps from probably to absolutely—typically occur. They do not take place primarily in classrooms, workplaces, or in conversations among mutually knowledgeable experts. They typically occur as expert testimony in courtrooms, where novice decisionmakers are charged with, among other things, weighing the value of that testimony. But the courtroom is a poor environment for elevating presumed probabilistic truths pertaining to forensic science evidence to scientific truths.

4. Policy

Even if forensic science individualization claims were supported by rigorous scientific testing, and even if Professor Kaye’s view that very low random match probabilities (if and when they were determined to exist in the various forensic sciences) were accepted as providing sufficient support for claims of individualization, good policy reasons counsel against permitting individualization testimony in criminal litigation.

First, in light of the long history of untested techniques, insufficient research, testimonial exaggerations, and fabricated findings in the forensic sciences, we should be hesitant about further elevating and legitimating unproven forensic science claims.

Second, given the adversarial nature of legal proceedings, the elevation of presumed truths or linguistic shortcuts to scientific ones unfairly privileges the offering party (which is usually the government).

Third, a seemingly harmless inferential leap from a very low probability to zero probability in the context of a criminal trial might have unintended consequences. One consequence is the suppression of uncertainty. When forensic scientists offer

---

65 See generally NRC Report, supra note 4.

66 Professor Kaye writes that “scientists have indicated that opinions of general uniqueness or uniqueness of a particular DNA type within some smaller region are or will soon become scientifically acceptable.” Kaye, supra note 5, at 1186 (footnotes omitted). If true, this phenomenon is being driven more by the exigencies of litigation than by the results of scientific research. As others have noted, this is less of a “scientific breakthrough,” as the idea was characterized when first announced, and more of a “semantic breakthrough.” William C. Thompson & Simon A. Cole, Psychological Aspects of Forensic Identification Evidence, in EXPERT PSYCHOLOGICAL TESTIMONY FOR THE COURTS 31, 45 (Mark Costanzo, Daniel Krauss & Kathy Pezdek eds., 2007).

an individualization conclusion, they signal to a factfinder that there is little point in weighing any evidence that militates against the expert’s conclusion. This could lead to serious errors by the factfinder.

The inordinate power of expert assertions is illustrated by a case in which a victim knew a suspect well and excluded him from a photo identification: she was sure he was not the man who raped her. Later, she was told that supposedly irrefutable scientific evidence pinpointed the man as her attacker. Induced to disbelieve her own personal knowledge, and given another opportunity to identify the suspect as being the rapist, she did so. At trial, the victim’s eyewitness identification was more dramatic and compelling than the “scientific” evidence, though it was a byproduct of the “scientific” evidence.68 The suspect, William O’Dell Harris, was convicted and sent to prison. Years later, it was learned that the “scientific” evidence had been fabricated.69 When the biological evidence was subjected to DNA testing, Harris was excluded as the rapist. After eight years in prison he was exonerated and released.70 If fabricated scientific evidence can cause a witness to disbelieve her own personal knowledge and accept a complete falsehood as true, surely factfinders are also susceptible to believing exaggerated testimony to the point of assuming that there is little uncertainty left to resolve in a case.

Fourth, research suggests that statements made by experts are given considerable deference by jurors and their impact is unlikely to be undone either through cross-examination or rebuttal witnesses.71

Fifth, when experts exaggerate the state of their science and their exaggerations find acceptance in the courtroom,

69 This was just one of myriad cases that the forensic scientist in the case, Fred Zain, was eventually found to have fabricated. Zain’s fabrications led the West Virginia Supreme Court to declare that, “as a matter of law, any testimonial or documentary evidence offered by Zain at any time in any criminal prosecution should be deemed invalid, unreliable, and inadmissible.” In re Investigation of the W. Va. State Police Crime Lab., Serology Div., 438 S.E.2d 501, 506 (W. Va. 1993) (adopting the findings of the Report of a special inquiry ordered by the Court).
70 The Innocence Project, Know the Cases, http://www.innocenceproject.org/know/ (last visited Feb. 8, 2010).
researchers have less incentive for conducting the basic and applied research needed to put these assertions to the test.\[72\] Thus, research on the frequency with which various characteristics occur and on the best ways to convey forensic science evidence may not even get off the ground.

Even if we reach a state where rigorous scientific support for individualization is available, there are many practical and policy reasons for not permitting the traditional forensic sciences to make the individualization leap.

IV. INDIVIDUALIZATION TESTIMONY: NOT HELPFUL

The 2009 National Research Council report on the forensic sciences called for more transparency and less exaggeration: “Forensic reports, and any courtroom testimony stemming from them, must include clear characterizations of the limitations of the analyses, including measures of uncertainty in reported results and associated estimated probabilities where possible.”\[73\] Fallacy offered a similar call. It argued that scientific foundations need to be improved, the application of those foundations to case-specific findings needs to be improved, and examiners’ personal views about the evidence need to be kept out of reports and testimony.

Given (a) the current lack of scientific support for claims related to individualization in the traditional forensic sciences, and (b) the likelihood that jurors will not meaningfully differentiate an examiner’s individualization opinions from a statement of scientific fact about individualization,” we suggest that forensic examiners should be barred from offering individualization opinions. Individualization opinions violate Federal Rule of Evidence 702, which requires that scientific opinion testimony, to be admissible, must “assist the trier of fact to understand the evidence or to determine a fact in issue”

\[72\] Cole, supra note 14, at 31-32 (discussing the “perverse incentives created by the current weak legal regime that permits extremely strong claims like ‘individualization’ without empirical support”). Professor Kaye appreciates this point as well. He notes that “a strong argument can be made” for excluding comments by an examiner related to why he thinks a match is probative of identity “to encourage more extensive research.” Kaye, supra note 5, at 1185 n.86.

\[73\] NRC Report, supra note 4, at 21-22; see also id. at 185.

and must be “based upon sufficient facts or data.”

Opinions about whether a marking has been individualized to its one and only possible source ordinarily are not based upon sufficient facts or data. Nor do they provide assistance to the trier of fact beyond that which can be gained from a less grandiose presentation of the forensic science findings, and a more candid presentation of their limitations. Instead, individualization testimony has considerable potential to mislead factfinders rather than to assist them. Though Federal Rule of Evidence 704 expressly permits the offer of an ultimate opinion (such as here, on identity), testimony admitted under Rule 704 still must pass the helpfulness requirements of Rule 702 and be based on adequate data to support the opinion. A forensic scientist’s opinion about source identification or individualization provides no more value to factfinders than what could be provided by more data-based statements, while having more potential to mislead.

In conclusion, forensic scientists should not be permitted to capitalize on the lack of supportive scientific data about either characteristic frequency or their own diagnostic reliability by going beyond what is known and what can be stated on good grounds. They should not be permitted to say, in effect, “trust me: that’s the source.” Real scientists don’t say “trust me.” They provide data.

---

75 Fed. R. Evid. 702.
76 “The abolition of the ultimate issue rule does not lower the bars so as to admit all opinions. Under Rules 701 and 702, opinions must be helpful to the trier of fact . . . .” Fed. R. Evid 704 advisory committee’s note.
77 Gary L. Wells, Naked Statistical Evidence of Liability: Is Subjective Probability Enough?, 62 J. Personality & Soc. Psychol. 739, 747 (1992) (finding, in Experiment 4, that both judges and jurors were far more likely to find liability when provided with expert testimony consisting of the relevant data plus the expert’s personal opinion that the defendant was the source, than when they were presented with all the same information but not the expert’s conclusory opinion).
The Courts, the NAS, and the Future of Forensic Science

Jennifer L. Mnookin

INTRODUCTION

On a recent flight, the person next to me on the crowded airplane began to chat with me. When I told her about what I researched and studied, she looked at me with a big grin. “I LOVE forensic science,” she said. “I watch CSI whenever I can. They can do such amazing things. It’s all so high tech—and incredibly accurate! It’s almost like magic, isn’t it?” She leaned in a bit closer and looked at me intently. “Tell me, is it like that in real life?”

I looked at her for a moment before answering. I felt a bit like the older child on the playground about to reveal to her younger friend that Santa Clause doesn’t really exist. I shook my head. “No, I wouldn’t say that CSI’s depiction is entirely realistic. In the real world, forensic science isn’t nearly so glossy. It isn’t nearly so speedy. And most important, it isn’t nearly so foolproof, either.”

“Really? That’s too bad,” she told me. She looked at me directly for a brief moment, shook her head, and then looked away. “Well, to tell you the truth, I think I’d rather just keep believing in the television version.” Figuring that reality was not going to be any match for CSI, I shrugged, and went back to the book I was reading.

In fact, that casual exchange on an airplane captures something quite important about the traditional forensic sciences, which find themselves at a crossroads. For many long-used types of forensic science, including fingerprint identification, firearms identification, handwriting identification, and toolmark

---

1 Professor of Law, UCLA School of Law. This is an expanded and revised version of the Belfer Lecture, presented at Brooklyn Law School on April 7, 2009, in honor of Professor Margaret Berger’s retirement. Professor Mnookin thanks Margaret Berger, Ed Chang, Itiel Dror, Jennifer Friedman, Jay Koehler, and D. Michael Risinger for helpful comments, conversations and suggestions. Many thanks to Forrest Havens for his helpful research assistance.
identification, experts’ claims about their field, the authority of their methodologies, and their own abilities have dramatically outstripped what has actually been established by persuasive research and careful study. Forensic scientists have regularly testified in court to matters that are, quite honestly, both less proven and less certain than they are claimed to be. They have overstated their degree of knowledge, underreported the chances of error, and suggested greater certainty than is warranted. More generally, many kinds of forensic science are not entirely based on the methods and approaches that we usually associate with validated research science. Their claims and the limits to their claims are not closely based on or constrained by the formal collection of data. Their empirical assertions are not grounded in careful research that has been subject to peer review and publication. There has been remarkably little formal validation of their methods. And there has been far too little study of how often forensic scientists might make mistakes, and when or why these possible errors are more likely to occur. Moreover, when academics attempt to do research on these questions, they have sometimes faced limited cooperation, or even downright resistance, from the forensic science community, because practitioners, managers, and laboratory directors (as well as police departments and prosecutors), are often wary of research not under their supervision or control.

For roughly the last decade, academic critics, and, occasionally, forensic scientists themselves, have argued that this state of affairs needs to change. If we cannot trust the

---

evidentiary inputs into our criminal justice system, we cannot trust the outputs either. Numerous well-publicized wrongful convictions have made the danger of error in our criminal justice system both more obvious and more salient. Recent research suggests that misleading and erroneous forensic science has been a significant contributing factor in many of the known wrongful convictions.\(^2\)


While the danger of erroneous conviction provides both a moral and practical perspective on why reliable and valid forensic science is so important, ordinary expert-evidence doctrine also mandates its validity as a prerequisite for admissibility. In 1993, the Supreme Court, in *Daubert v. Merrell Dow*, made clear that judges have a gatekeeping responsibility with respect to expert evidence. In the federal courts and in those states that have embraced *Daubert*, expert evidence needs to be sufficiently reliable—meaning, more or less, scientifically valid—in order to be legitimately admissible in court. As a matter of formal evidence doctrine, then, forensic science evidence should only be permitted if it meets *Daubert*’s requirements. While *Daubert* envisioned the judicial gatekeeper’s inquiry into reliability as “flexible,” and therefore did not set up any absolute criteria for determining the validity (and hence admissibility) of expert testimony, the majority opinion did provide some important guidelines for trial court judges. Specifically, *Daubert* invites courts to look at whether the evidence or technique in question has been tested adequately, whether it has a known error rate; whether it has been subject to peer review; and whether it is generally accepted by the relevant scientific community. At this point, numerous *Daubert* challenges have been made to many kinds of forensic science, from fingerprint evidence, to ballistics analysis, to handwriting examination. But with a small

---

5 Justice Blackmun stated that the “inquiry” into the admissibility of expert evidence is “a flexible one,” whose “overarching subject is the scientific validity—and thus the evidentiary relevance and reliability—of the principles that underlie a proposed submission.” *Daubert*, 509 U.S. at 594-95.
7 For examples of courts permitting forensic science under *Daubert* or the state equivalent, see United States v. Havvard, 260 F.3d 597 (7th Cir. 2001) (finding that latent fingerprint identification methods satisfied the standards of reliability set forth by *Daubert*); United States v. Mitchell, 145 F.3d 572, 574, 579-80 (3d Cir. 1998); *see also* United States v. Ford, 481 F.3d 215, 222 (2d Cir. 2007) (finding shoeprint analysis to satisfy the *Daubert* standard); United States v. Mitchell, 365 F.3d 215, 233, 250 (3d Cir. 2004); United States v. Taylor, 663 F. Supp. 2d 1170, 1180 (D.N.M. 2009) (finding firearm identification evidence to satisfy the *Daubert* standard); *State v. Foreman*, 954 A.2d 135 (Conn. 2008); *United States v. Mahone*, 453 F.3d 68, 79 (1st Cir. 2006); *United States v. Prime*, 363 F.3d 1028, 1034 (9th Cir. 2004) (finding that handwriting analysis satisfies the *Daubert* standard); United States v. Crisp, 324 F.3d 261, 271 (4th Cir. 2003); *Howard v. State*, 853 So. 2d 781, 796 (Miss. 2003) (finding that bite mark analysis satisfies the Daubert standard). For a list of Daubert challenges to fingerprint evidence (but current only through 2005), see http://onin.com/fp/daubert_links.html (last visited Feb. 26, 2010). For the argument that courts have been less intense in their Daubert scrutiny in criminal cases than in civil, see, e.g., D.
number of exceptions, courts have continued to permit these kinds of evidence without limit.

The truth of the matter is that for the last decade, both judges and the forensic science community have chosen to behave rather like my acquaintance on the airplane. They prefer to play a kind of make-believe; they prefer to believe in the television version.

In what follows, I aim to do three things. First, in Part I, I will provide a brief overview of the present state of affairs within forensic science, focusing on latent fingerprint evidence and the concerns that have emerged regarding the adequacy of its research basis. Second, in Part II, I will discuss a recent and significant report issued in February 2009 by the National Academy of Sciences regarding the needs of the forensic science community. I will suggest that this report, though impressive in many ways, gave too little attention to how the courts ought to handle the admissibility of pattern identification evidence. I wholeheartedly agree with this report’s assertion that it is imperative that we create substantially more funding for research and government oversight and regulation of forensic science. But this alone will not be enough, nor is it likely even to happen at all, unless courts also begin to take their responsibilities in this area more seriously. In Part III, I will look closely at how the courts have confronted (or, more accurately, mostly avoided confronting) the present problems relating to forensic science evidence in recent years. I will describe the approaches, mechanisms and machinations by which numerous courts have failed to treat their responsibilities to assess the validity of forensic science


There have been a small handful of cases that have restricted forensic science evidence, at least to some extent; this issue will be discussed in detail infra Part III.C. See, e.g., Taylor, 663 F. Supp. 2d at 1180; United States v. Green, 405 F. Supp. 2d 104, 124 (D. Mass. 2005); United States v. Hines, 55 F. Supp. 2d 62, 73 (D. Mass. 2002); United States v. Llera Plaza (I), 179 F. Supp. 2d 492 (E.D. Penn. 2002) (overruled by Llera Plaza II, 188 F. Supp. 2d 549 (E.D. Penn. 2002)); Maryland v. Rose, No. K06-0545 (MD Cir. Ct. Oct. 19, 2007). Expert evidence in handwriting identification has been scrutinized more carefully by courts than have the other forms of pattern identification evidence, and this scrutiny has sometimes led judges not merely to limit but to exclude it altogether. For a thoughtful account of the current case law in this area, see D. Michael Risinger, Handwriting Identification, in FAIGMAN, MODERN SCIENTIFIC EVIDENCE, supra note 1; D. Michael Risinger, Cases Involving the Reliability of Handwriting Identification Since the Decision in Daubert, 43 TULSA L. REV. 477 (2008) [hereinafter Risinger, Cases Involving].

evidence with adequate care, and I will also describe the approaches taken by the few courts who have addressed the issues seriously. This section also offers the first detailed scholarly analysis of an approach taken by a handful of thoughtful jurists with respect to forensic pattern evidence, in which they continue to admit the evidence but only in a weakened, limited form—specifically, they permit the expert to describe similarities and differences between exemplars, while excluding the experts’ ultimate conclusions about identification. This section describes why that approach, though superficially quite appealing, is far less conceptually coherent than it appears, though I recognize that it may nonetheless remain an attractive, pragmatic, stop-gap measure for courts wrestling with these difficult issues. Finally, in Part IV, I describe what a serious judicial examination of forensic pattern identification evidence ought to entail, what questions judges should focus on under *Daubert v. Merrell Dow*, and what demands they should make as a prerequisite to admissibility in court. I suggest that outright exclusion may, in some cases, indeed be warranted, and should certainly, along with more modest measures, be part of the available judicial toolkit.

A few brief preliminaries are needed to provide context and background. First, it is important to recognize that these concerns are not merely abstract or theoretical. In fact, a recent study of wrongful convictions found that flaws with forensic science—including interpretive errors, overstated testimony, and inaccuracies—were present in a whopping 60% of the cases studied. To be sure, many of the forensic science errors found

---

10 Garrett & Neufeld, *supra* note 3, at 19 n.62. In the case of microscopic hair analysis, it is important to distinguish those cases in which forensic scientists engaged in misstatement, overstatement, and unjustified assertions from those cases in which the inherently limited sensitivity of the technique means that in retrospect, the jury likely made an incorrect inference from the forensic evidence. Practitioners of microscopic hair identification have never formally claimed an ability to identify an individual from a hair; at most, a hair can be said to be “consistent” with a source, which means only that it is a member of a class of hairs that could possibly have come from that source, not that the hair necessarily came from that source. In this sense, it is more like blood typing—e.g., “this blood is type A, and so is the defendant, so the defendant is not excluded from the group of people who are possible sources for this blood” than it is like DNA profiling or fingerprinting. In my opinion, if a microscopic hair examiner provided only this kind of “class” evidence, including the defendant (or other relevant person) as a possible source among other possible sources, it ought not to be considered a forensic error, even if subsequent evidence shows the defendant not to be the perpetrator. It is as if the (innocent) defendant, the blood sample, and the actual perpetrator all had Type A blood—for the serologist to have testified that the defendant’s blood type matched the crime scene was not erroneous, though to be sure, the inference of guilt the jury drew from that fact combined with the other evidence
in the study related to techniques that play a less significant role as legal evidence than they once did, such as microscopic hair examination (which now most frequently functions as an adjunct to mitochondrial DNA testing of hairs instead of standing alone as it used to do).”

However, to intimate that therefore these past errors and mischaracterizations no longer matter is to miss a critical point. While mitochondrial DNA testing can now often operate as a check on microscopic hair analysis, the broader forensic science ‘culture’ which made these earlier instances of error possible remains very much intact. This forensic science culture—a culture in which claims derived from experience are often accepted as a substitute for data; a culture in which interpretations are often framed in absolute terms rather than in more limited or modest language; a culture in which potentially biasing information is not systematically kept from the forensic examiner; and a culture in which institutionally cozy relationships between detectives, forensic analysts, and prosecutors may encourage unconscious partisanship—remains very much the norm within forensic science laboratories today. It is, in the end, this culture that needs to change; new and improved forensic techniques will not, by themselves, provide an adequate solution.

Second, even though I certainly do wish to criticize the current paltry research basis of forensic science and the courts’ response to that reality, I also want to emphasize several important caveats. Forensic professionals are, for the most part, just that—hard-working, dedicated, and trying their best, often with quite inadequate funding. Though there are, unfortunately, more than a handful of known instances of forensic fraud, these incidents are the exception and not the

\[\text{was erroneous. By contrast, when, as was all too often the case, microscopic hair analysts claimed or intimated that the defendant was the source, or provided fictitious frequency estimates for hair types, or claimed the hairs “matched” without making clear that numerous other peoples’ hair in any given population would also match, this would, in my view, certainly count as forensic error. For examples of forensic error, see generally id. I thank Barry Scheck for useful conversations on this point that clarified my thinking, even if we continue to disagree on some definitional points.}\]

rule. Most forensic scientists are both diligent and honest. My criticisms of the field should not be heard as criticisms of all those who pursue it. Furthermore, the lack of an adequate research basis for the claims of the pattern identification sciences is not the fault of practicing rank-and-file forensic scientists themselves, most of whom have neither the training nor the background to pursue such research effectively even if they wanted to do so. These practitioners should—and indeed must—become, to some extent, the subjects of research inquiry, but they are not to be faulted for failing to become empirical researchers themselves.

Third and finally, even though I strongly believe that forensic science needs to be placed on a more secure research foundation, I do not want to suggest, or to be heard to suggest, that it is therefore of no value. In fact, with most of the forensic sciences, my strong suspicion is that when we do finally insist on pursuing the necessary research, we will find that many kinds of forensic evidence presently in use turn out to be extremely probative, and very much worth hearing in court. In many fields, my prediction would also be that the error rate, even with the methodologies in use at present, will turn out to be tolerably low for a wide array of pattern identification tasks, though I am far from confident this will turn out to be true in all fields, or in all situations. But for now, these expectations are mere speculation—the critical point is that it is time to pursue research that will help us find out for sure. What we have, at present, is no more and no less than an absence of adequate evidence. This lack of evidence does not in and of itself prove the inadequacy of the methods used by forensic scientists. But we can, should, and therefore must, seek better evidence establishing the validity of these techniques, methodologies, and conclusions, so that we can have greater legitimate confidence in the forensic sciences we use in court, and thus better understand their possible limits and weaknesses as well.

I. FINGERPRINT EVIDENCE AND THE LIMITS OF OUR KNOWLEDGE

In what follows, I focus primarily on latent fingerprint identification, but it is important to realize that I could tell an extremely similar tale about a variety of other kinds of forensic science, including firearms identification, handwriting identification, bitemark identification, toolmark identification, and the like. Fingerprint evidence is, in all likelihood, both more probative and less error-prone than some other kinds of forensic identification evidence, and it has a long and extremely substantial courtroom use. It therefore provides an especially good focus, for if the problems I am describing exist within this forensic domain, they are likely to be equally or more acute in other areas of pattern identification.

Fingerprint evidence was first used in the American courtroom nearly a century ago in 1911, and for most of its history it has been seen as the “gold standard” of forensic science. In recent years, however, whatever metal out of which this evidentiary standard was made has rather noticeably begun to tarnish. \footnote{For the history of fingerprint identification and its legal use, see generally Colin Beaven, Fingerprints: The Origins of Crime Detection and the Murder Case that Launched Forensic Science (2001); Cole, Suspect Identities, supra note 1; Simon A. Cole, Witnessing Identification: Latent Fingerprint Evidence and Expert Knowledge, 28 Soc. Stud. Sci. 687 (1998); Mnookin, Fingerprint Evidence, supra note 1.}

The basic approach taken by latent fingerprint experts involves what they call ACE-V. This acronym stands for analysis, comparison, evaluation and verification. \footnote{For a description of each stage, see, e.g., Michell Triplett’s Fingerprint Dictionary, available at http://www.nwlean.net/fprints/a.htm.} First, in the analysis step, the examiner looks closely at the latent print associated with the crime at issue, and decides whether there is enough useful information contained in the image that it is “of value” for further examination. \footnote{Id.} If so, the examiner then looks carefully at the various minutiae that he or she sees in the image, and, depending on local practices and the apparent difficulty of the print, typically marks up the print and documents the minutiae she observes. \footnote{See, e.g., Herman Bergman & Arie Zeelenberg, Fingerprint Matching, Manual, in Encyclopedia of Biometrics (Anil K. Jain & Stan Z. Li., eds), at 502-04 (2009).} Second, in the comparison stage, the expert compares the latent print to a
particular source print, noting both observed similarities and differences.\textsuperscript{17} Third, the analyst evaluates these similarities and differences, and reaches one of three, and only three, conclusions: identification, exclusion, or inconclusive.\textsuperscript{18} Note that these are the only permissible options available to a latent fingerprint expert—a match, a non-match, or a conclusion of “I don’t know.” “Maybe,” “possibly,” and “probably,” are not determinations presently permitted to fingerprint examiners under their professional rules and norms.\textsuperscript{19} Finally, in the verification step, if the first examiner has determined that the prints match, a second examiner takes the prints and goes through the same process to re-analyze them. In most laboratories, this step is conducted by an examiner who is informed of the original examiner’s conclusion before undertaking his or her own analysis.\textsuperscript{20} This verifying examiner typically recognizes both that (a) he or she is verifying a conclusion already reached by someone else; and (b) that the conclusion already reached is that the prints do match.

Latent fingerprint examiners regularly claim that ACE-V is a version of the scientific method and assert that it offers a reliable methodology that establishes that fingerprint evidence is indeed a valid science.\textsuperscript{21} Many courts have agreed that ACE-V passes muster under Daubert.\textsuperscript{22}

\textsuperscript{17} The comparison print may have been to a known suspect, or to a non-suspect known to have been in the relevant location. Or the latent print may be submitted for an Automated Fingerprint Identification System (AFIS) database search, which compares the latent to a large database (depending on the particular database, often many millions of images) and returns a set of possible prints for human comparison. For some of the difficulties with interpretation of match thresholds in AFIS searches, see Itiel Dror & Jennifer Mnookin, The Use of Technology in Human Expert Domains: Challenges and Risks Arising from the Use of Automated Fingerprint Identification Systems in Forensic Science, 10 LAW, PROBABILITY & RISK (forthcoming, 2010).

\textsuperscript{18} For descriptions of these permitted conclusions, see SWGFAST (Scientific Working Group on Friction Ridge Analysis, Study, and Technology), Standard for Conclusions, http://www.swgfast.org/Standards_for_conclusions_ver_1.0.pdf.

\textsuperscript{19} Note that at its 2009 annual meeting, the International Association of Identification (IAI), the professional organization of fingerprint examiners, considered eliminating the longstanding professional restriction on testifying in probabilistic terms. However, the motion was the subject of significant contention and was tabled for further study.

\textsuperscript{20} The FBI has begun to conduct some verifications “blind,” meaning that the verifier does not recognize that he or she is verifying a conclusion reached by another examiner. See Robert B. Stacey, A Report on the Erroneous Fingerprint Individualization in the Madrid Train Bombing Case, 54 J. FORENSIC IDENTIFICATION 706, 715 (2004).

In fact, ACE-V’s relationship to the scientific method is tenuous at best: as a methodology, it amounts, more or less, to having two different examiners look carefully at a set of fingerprints. To be sure, the “scientific method” is itself a complicated and capacious idea, not altogether easily or adequately defined. But however we might define the critical characteristics of the scientific method, it surely amounts to more than simply careful, semi-structured observation. At root, ACE-V in its current incarnation amounts to no more and no less than a set of procedures to describe the careful comparison of a latent print with a potential source print by an initial examiner and a subsequent verifier. While careful observation and the recording of one’s observations may be a necessary part of many scientific practices, careful observation in and of itself cannot be meaningfully said to constitute a method. Moreover, the simple act of labeling this process of careful observation as a methodology does not make it into one. Nor does bestowing upon it the label “scientific” tell us, through the moniker, anything about its likely validity or error rate.\(^2\)

The basic difficulty is that ACE-V is too general in conception and scope to provide much in the way of guidance or constraint for those who practice it. The devil is in the details—what constitutes analysis? How exactly does a competent comparison take place? When are apparent similarities misleading, and when might apparent differences be attributed to something other than the two prints deriving from different sources? ACE-V, as a methodology, does not help answer any of these critical methodological questions, because its categories are too general and insufficiently substantive.\(^3\)

It is as if one were to describe the methodology for fixing a car by the acronym DACT—Diagnose, Acquire, Conduct, and Test. We could describe the DACT car-repair methodology as described ACE-V as corresponding to the scientific method. See e.g., Mary Beeton, **Friction Ridge Identification Process—Proposed Scientific Methodology, THE DETAIL, Feb. 18, 2002, available at** [http://www.clpex.com/Articles/TheDetail/1-99/TheDetail28.htm](http://www.clpex.com/Articles/TheDetail/1-99/TheDetail28.htm).

---


\(^3\) For the argument that the point ought not to be whether fingerprint evidence is or is not scientific, but rather, how to improve it, see Itiel E. Dror, **How Can Francis Bacon Help Forensic Science? The Four Idols of Human Biases**, 50 JURIMETRICS J. 93 (2009).

\(^4\) See generally Haber & Haber, supra note 1.
follows: (1) diagnose the car’s problem, (2) acquire the necessary parts for the repair, (3) conduct the repair, and (4) test to verify that the repair fixed the problem. Whether or not such a car-repair methodology actually works, or how well it works, would depend entirely on the content given to these very broad categories in specific instances. If in fact, someone diagnosed the car’s problem correctly, located the appropriate parts, and conducted the repair properly, the methodology would work. But if the mechanic misdiagnosed the difficulty, acquired the wrong parts, or made an error when conducting the repair, the repair would fail, even though he or she had, in some sense, followed the methodology. Now, in light of the failed repair effort, a defender of ACE-V (and DACT) might suggest that the mechanic had not in fact followed DACT correctly, because he or she misdiagnosed the problem, made an error in the repair, or made some other mistake in application. The DACT defender might even argue that the mechanic’s failure to fix the car established that she failed to follow DACT; that following DACT necessitates doing the steps correctly, not just endeavoring to follow them. But that response would render DACT (or, analogously, ACE-V), in some sense, merely tautological. The method does not describe with any specificity how to complete its requirements correctly. It is therefore illegitimate to argue that the method has not been followed simply because the desired outcome did not occur, precisely because the method itself underspecifies what is required. DACT itself does not explain how to diagnose, or what constitutes a sufficient repair, just as ACE-V does not explain how to analyze or compare (beyond calling for careful looking at a target portion of each print), or what constitutes a sufficient evaluation. To be sure, ACE-V might be a useful description of the basic steps a fingerprint examiner takes in order to conduct his or her examination, but that does not make it a very useful description of a methodology, much less a so-called “scientific method.”

25 I am presently a member of an NIST/NIJ working group on Human Factors in Latent Fingerprint Examination. As part of this working group’s efforts, fingerprint examiners put together a process map describing the ACE-V process in more careful detail than had ever previously occurred. While the process map is useful in many ways, it also makes my point here clear: although the process map delineates numerous steps, and diagrams in detail the steps necessary for conducting a comparison, the actual interpretive content of terms like “sufficiency,” “identification,” etc., are never actually specified, because at present, the fingerprint community does not have shared definitions of these concepts.
Surely, one would think, ACE-V in practice must amount to more than I am suggesting? Latent print examiners do have norms about what kinds of print ridge detail and minutiae they ought to be looking at, and examiners are trained to search both for relevant minutiae and to assess their contextual relationship and position on a fingerprint.\textsuperscript{26} And latent print examiners do discuss with one another, informally, their personal notions regarding sufficiency, or the virtues and limitations of different categories of print information.\textsuperscript{27}

While individual examiners or even sometimes laboratories may develop working rules of thumb about the quantity of similarity required, latent fingerprint examination as a field lacks any formalized specifications about what is required in order to declare a match. There is no required minimum number of points of resemblance or minimum number of total print features, nor any required quantum of any specific kind of ridge detail.\textsuperscript{28} Instead, examiners decide for themselves, based on their training and experience, how much similarity is sufficient to declare a match. Moreover, when examiners look at a print, they may not even be focusing on the same features. Two fingerprint analysts will often focus on different minutiae in their examination of the same print; indeed, sometimes the same examiner, when given the same print at a different time, will focus on different minutiae than

\textsuperscript{26} Level 1 detail describes the major pattern of the print, whether it is, say, a tented arch or a loop. Level 2 detail, which constitutes the main focus for comparison and evaluation for most examiners, refers to the ridge quantities, details, and characteristics, such as whether and where a friction ridge bifurcates, or terminates, or develops a spur. Level 3 detail refers to sub-ridge detail, for example sweat pores that may be visible in an image.

\textsuperscript{27} Fingerprint examiners, though, do not all agree that it is useful to divide the field into precisely three different levels. See supra note 25. Though the terms are often used in the United States, examiners in the UK do not typically focus on these categories, and there are diverging opinions on how to define and how to make use of Level 3 detail, for example. See the definitions of Level 1, 2, and 3 detail in Michelle Triplette’s Fingerprint Dictionary, supra note 14 (available at http://www.nwlean.net/fprints/l.htm); see generally David R. Ashbaugh, Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology (1999).

\textsuperscript{28} Indeed, the IAI (International Association of Identification) resolved in 1973 that there was no scientific justification for having a specified minimum number of points of similarity in order to declare that two prints matched. See Report of the Standardization Committee of the International Ass’n for Identification, Identification News, Aug. 1, 1973, available at http://www.latent-prints.com/images/IAI%201973%20Resolution.pdf. For further discussions of this point, see generally Mnookin, Confessions of a Fingerprinting Moderate, supra note 1; Simon A. Cole, What Counts for Identity? The Historical Origins of the Methodology of Latent Fingerprint Identification, 12 Sci. in Context 139 (1999) [hereinafter Cole, What Counts for Identity?].
he or she did the first time.\textsuperscript{29} The judgment is fundamentally a subjective one, not based on any formalized measures of either quantity or sufficiency.\textsuperscript{30}

Additionally, latent fingerprint examiners do not generally employ any statistical information or models in the ordinary ACE-V process. The field presently does not have or make use of robust statistically-based data about the frequency of different friction ridge characteristics. Analysts do not make regular or structured use in their comparisons of empirical studies showing how common or how rare different fingerprint details might be. They do not presently make use of any statistically-validated standards to justify how many identifying characteristics must be the same on two prints in order to warrant a finding that they match. Nor do they employ a probabilistic approach to determining the likelihood that a print selected at random would have that quantum of similarity, akin to the use of “random-match probability” in DNA identification. Although significant strides are being made toward developing these kinds of information, technical obstacles still limit the ability to develop a satisfactory statistical measure of the frequency of various ridge characteristics.\textsuperscript{31} As of now, there simply is no well-accepted, fully-specified statistical model that is available for latent fingerprint examiners to employ.

The list of difficulties continues. A fundamental tenet of latent fingerprint analysis is the “one discrepancy rule”—if there is even one genuine discrepancy between the latent print and a potential source print, then the two prints cannot have come from the same source.\textsuperscript{32} This, however, invites the critical question of how to decide what constitutes a discrepancy, as

\textsuperscript{29} Itiel E. Dror et al., \textit{Cognitive Issues in Fingerprint Analysis: Inter- and Intra-Consistency and the Effect of a Target Comparison} (unpublished article, under review at \textsc{Forensic Sci. Int'l}) (on file with author).

\textsuperscript{30} See generally id.; Mnookin, \textit{Confessions of a Fingerprinting Moderate}, \textit{supra} note 1; Mnookin, \textit{Fingerprint Evidence}, \textit{supra} note 1; Saks, \textit{Merlin and Solomon}, \textit{supra} note 1; Cole, \textit{What Counts for Identity?}, \textit{supra} note 28.

\textsuperscript{31} One quite promising approach is being developed by a group of European researchers. See, e.g., Cedric Neumann et al., \textit{Computation of Likelihood Ratios in Fingerprint Identification for Configurations of Any Number of Minutiae}, 52 J. of \textsc{Forensic Sci.} 54, (2007); J.S. Buckleton, C.M. Triggs & C. Champod, \textit{An Extended Likelihood Ratio Framework for Interpreting Evidence}, 46 \textsc{Sci. and Just.} 69 (2006). While a version of software implementing this team’s approach to providing probabilistic likelihood ratios is likely to be available soon, its adequacy and validity is not yet fully established.

opposed to a dissimilarity that can legitimately be explained in some other way. The problem is that no two print impressions are ever truly identical—every single impression from a print is distinct from every other impression of a print, different to some extent even from those that came from the same source. A print image can be affected by the pressure with which it was left, the surface on which it was made, the processes by which it was lifted, and many other factors. The question when comparing two prints, then, is not whether they are truly “identical”—for they will never be truly identical—but rather, whether they are sufficiently similar to each another to permit the conclusion that they came from the same source. The examiner needs to determine whether apparent differences are true dissimilarities, or instead, merely artifacts that ought not to be deemed meaningful. Unfortunately, latent fingerprint examiners lack any formalized criteria for determining when a difference between two prints is genuinely a dissimilarity, or when it might appropriately be explained in other ways. At root, this is again a matter of subjective judgment by the trained examiner.

Note, however, that the fact that these judgments are subjective does not necessarily imply that they are incorrect or unreliable. If I were to look at many different photographs of my sister, no two images of her would be identical. And yet, my judgment of whether any given photograph was an image of my sister or actually an image of someone else bearing a certain degree of resemblance to my sister would, I would wager, have a high probability of being correct. I would posit that my ability to identify images containing my sister, and to avoid misidentifying images of other people as my sister would be quite high—notwithstanding my lack of formal criteria for doing so. The absence of formal, validated standards for making such identification of my sister does not mean that I lack all relevant knowledge. My experience of many years of seeing my sister in a great variety of contexts would indeed likely help me with the identification tasks.

However, I also suspect that my ability to identify my sister in photographs would be strong but not perfect. In some images, she might be too far away, or too blurry, or someone else might bear such a strong resemblance to her, that despite my life-long knowledge of her from every angle, I might nonetheless mistake the other person for my sister. Or conversely, I might fail to recognize that some picture truly did show an image of my sister. Of course, to determine how often I
was right or wrong, we would also want to make sure we had a good method by which to determine “ground truth,” whether or not the photograph truly was of my sister.

The purpose of this analogy is to suggest three points that apply as much to fingerprint identifications as to my hypothetical efforts to identify my sister. First, I want to suggest that experience can be a legitimate basis for knowledge. Second, I want to suggest that knowledge need not necessarily be formalized to count as legitimate or valid. However, and this is the third point, if we wanted to find out just how good my ability to recognize my sister in photographs really was, we would need to depend on something that went beyond my say-so. We would not want simply to take my word for it when I said I was good at the task. We would not want to take the simple fact of my extensive experience looking at my sister as proof of my identification talents. Nor would we want to blindly accept my opinion that particular photos actually were or were not of my sister.

Instead of taking my say-so, my experience, and my conclusions as proof of my accuracy, we should carefully test my actual proficiency at the tasks. We would need to investigate empirically just how well I did identify my sister; in what array of circumstances I succeeded; and when and how often the task proved beyond my capacity. Indeed, as I will argue below, an equivalent focus on serious, careful proficiency testing of practitioners is precisely what we ought to demand in the realm of forensic science as well. Just as we would want proficiency testing to verify my claimed experience-based ability to identify my sister, so we also ought to require significant proficiency of fingerprint examiners and other pattern identification analysts. 33 And just as we ought not to simply take my assertions about my conclusions’ accuracy as proof of actual accuracy, we ought not to take fingerprint examiners’ experience-based assertions of accuracy as proof of accuracy either. 34

33 See infra note 157 and accompanying text (discussing proficiency tests for fingerprinting).
34 One important issue is feedback: if I do make mistakes in identifying my sister, am I likely to know about them? Do I receive feedback on my accuracy that could permit me to learn from my errors? If not, we should be especially skeptical of my assertions of accuracy based on experience, for if my experience is not likely to provide me with information about when and where I went wrong, I may not be able to learn all that much from it. This is clearly a concern with the pattern identification sciences as well, in which the opportunities for learning from feedback are quite limited. (The situation may be even worse than one in which there is no feedback; the trial process
Returning now to latent fingerprint examination, there are two additional difficulties with the current state of knowledge and practice. First, fingerprint experts claim to be able to individualize—to connect a given print to a unique source. When they declare a match, they assert that two prints come from a common source to the exclusion of all other possible sources in the world. This is an astonishingly strong claim. A latent fingerprint examiner who individualizes is saying that he or she can connect this print to one particular finger of one single person, out of everyone in the world, everyone who has ever lived or will ever live. But there is quite simply a lack of empirical evidence establishing that they can actually do what they claim. To be sure, both experience and some empirical research does suggest that fingerprints are highly varied. But even if fingerprints themselves are unique, this does not necessarily mean that experts can make unique identifications from partial latent prints, using their methods and expertise. Evidence of uniqueness does not itself directly support the experts’ claims that they are able to individualize.

may provide erroneous or misleading feedback in the case of forensic errors, precisely because forensic science evidence is often considered to be so strong. For example, if an expert makes an erroneous fingerprint identification that leads to a conviction or a guilty plea, that could be seen by the expert as confirming the correctness of the call, when in fact the erroneous identification may have substantially produced the conviction."


NAS Report, supra note 9; Mnookin, Confessions of a Fingerprinting Moderate, supra note 1; David A. Stoney, What Made Us Ever Think We Could Individualize Using Statistics?, 31 J. FORENSIC SCI. SOCY 197, 197 (1991); Saks & Koehler, Individualization Fallacy, supra note 1.

See, e.g., an unpublished study cited and described in United States v. Mitchell, 365 F.3d 215 (3d Cir. 2004) as the 50k study, designed to show the tremendous variation in fingerprints. But for a strong critique of this study, see David H. Kaye, Questioning a Courtroom Proof of the Uniqueness of Fingerprints, 71 INT'L STAT. REV. 521 (2003).

See generally Cole, Grandfathering Evidence, supra note 1; Mnookin, Confessions of a Fingerprinting Moderate, supra note 1. Think again of my sister in the photograph, discussed supra notes 33-34 and accompanying text. That every person really is unique does not mean that every photographic image will be a sufficiently clear depiction to permit a unique identification, or that my sister-identification talents will be strong enough to succeed in all circumstances. Even if every person (or every fingerprint) is indeed unique, the question is whether the image and the methods used to analyze it are capable of discerning that degree of difference.
Nonetheless, experts claim this ability as well as frequently asserting 100% certainty in their own conclusions. In fact, they are prohibited under their professional norms from making probabilistic judgments, and are subject to possible sanction if they do not follow this professional rule. They are told and taught that they must either be absolutely certain, or reach no conclusion at all. There are no shades of grey permitted—withstanding that fundamentally, fingerprint matching ought to be thought of as a probabilistic inquiry.

How often do fingerprint examiners make mistakes? Well, in court, until quite recently, experts frequently testified that their technique had a “zero error rate.” Some examiners tried to divide their analysis of error rate into two parts—the error rate of the technique itself, and the error rate of the humans who use it. They acknowledged that it was possible for a human to make a mistake, though they asserted (again without any significant published research) that errors are exceedingly rare. They claimed that the technique, if used properly, is perfect and error-free. Mistakes result only when humans misapply it.

This notion of an error rate of zero is exceedingly unscientific. It borders on the meaningless, and is a far cry from how scientists typically think about error rates. Nothing is truly perfect—no human endeavor has an error rate of zero. Moreover, the distinction between the error rate of the technique and the error rate of the humans who use it is, frankly, nonsensical with regard to fingerprint identification. The human beings engaging in ACE-V are the technique. The appropriate question is the error rate in practice, not an in-the-clouds theoretical error rate that postulates perfect human

---

See generally SWGFAST Standards for Conclusion, supra note 18; supra text accompanying note 19.

Mnookin, Confessions of a Fingerprinting Moderate, supra note 1; Cole, More than Zero, supra note 35, at 992; Cole, Forensics Without Uniqueness, supra note 35, at 235. However, there is ongoing discussion in the fingerprint community about whether to modify or eliminate this rule. See supra text accompanying note 19; Stacey, supra note 20, at 715.

The FBI’s Steven Meagher was perhaps the best known example of an expert who regularly testified in this vein. See, e.g., Mitchell, 365 F.3d at 222-26 (discussing Meagher’s testimony); see also, e.g., United States v. Mahone, 453 F.3d 68 (Me. 2006) (zero error rate argued in footwear identification case, drawing on latent fingerprint for support and by analogy); Havvard, 260 F.3d 597; Llera Plaza (I), 179 F. Supp. 2d 492.

See, e.g., Mitchell, 365 F.3d at 222-26 (discussing Meagher’s testimony); Llera Plaza (I), 179 F. Supp. 2d 492; see also Mnookin, Confessions of a Fingerprinting Moderate, supra note 1; Cole, More than Zero, supra note 35, at 1050.
beings and then concludes that so long as these perfect human beings make no mistakes, the error rate is zero. We could just as easily say that in theory, eyewitness identification has an error rate of zero because faces are in fact different—notwithstanding the fact that in practice, eyewitness identification errors are distressingly common. This claim of an error rate of zero is an example of how the rhetoric of forensic sciences is often shaped for courtroom use rather than derived from valid scientific testing designed to produce the most accurate possible information.

If the actual error rate is not zero, then how often do fingerprint examiners make mistakes? The truth is that we really do not know. Although fingerprint examiners may, at times, undergo proficiency tests, these exams have for the most part been extremely easy, far easier than the kinds of challenges that can be faced in actual casework. A fingerprint examiner from Scotland Yard once, under oath in court, referred to the FBI’s proficiency tests as a “joke” because of how ridiculously easy they were. Furthermore, in some laboratories, examiners take their proficiency tests in groups rather than individually. In addition, proficiency tests are generally not conducted blind, as part of what appears to the examiner to be ordinary casework. Rather, examiners are usually aware they are being tested, and may therefore, consciously or unconsciously use a different degree of care than usual. Thus, in their current form, proficiency tests might be a check on gross individual incompetence, but they certainly provide nothing more, and given that the examinations may be done collectively and not blind, they may not even necessarily provide that.

We do know that errors sometimes occur, though it is impossible on the basis of what we presently know to attempt

---

44 See Mnookin, Confessions of a Fingerprinting Moderate, supra note 1, at 137; see also Mnookin, Scripting Expertise, supra note 1, at 1727.
45 See Mnookin, Confessions of a Fingerprinting Moderate, supra note 1, at 135-36; Koehler, Proficiency Tests, supra note 1, at 1092; Cole, More than Zero, supra note 35, at 1032.
46 Llera Plaza (II), 188 F. Supp. 2d 549.
47 Koehler, Proficiency Tests, supra note 1, at 1092; Mnookin, Confessions of a Fingerprinting Moderate, supra note 1, at 136.
48 Koehler, Proficiency Tests, supra note 1, at 1092.
to quantify their frequency.\(^{49}\) One particular fingerprint error—perhaps the monster of all fingerprint errors, the most high-profile, embarrassing fingerprint mistake in recent history, at least here in the United States—has contributed to shaping and framing the discourse surrounding latent fingerprint identification.\(^{50}\) This mistake was sufficiently public, serious, and embarrassing that it led to a substantial inquiry into its causes; more generally, it made the fingerprint community—and the legal community—recognize that fingerprint errors were not simply a matter of incompetence or an issue of purely academic concern.\(^{51}\) I am referring, of course, to the mistaken identification of Brandon Mayfield, an attorney from Portland, Oregon who was held as a material witness in relation to the 2004 Madrid train bombing. The only evidentiary basis for suspecting his involvement was an alleged fingerprint match. Mayfield’s print had been one of the possible source prints suggested by a computer database search using an AFIS (Automated Fingerprint Identification System). Mayfield’s print appeared fourth down on the computer-generated list of suggestions—and according to the FBI, his print was a definite match.\(^{52}\)

Mayfield insisted that the identification had to be a mistake. He told authorities he had never set foot in Spain, had remained entirely in the United States during the relevant period, and indeed, lacked a passport. But three separate fingerprint examiners at the FBI, including two of the most respected senior examiners in the office, all concluded that the match was 100% certain. Even an independent, court-appointed expert confirmed the match as well.\(^{53}\)

\(^{49}\) For an effort to describe the array of known errors, see Cole, More than Zero, suprnote 35.

\(^{50}\) The ongoing saga in the United Kingdom relating to Shirlie McKie might offer the Mayfield case some competition.

\(^{51}\) This increased awareness, openness, and a certain increased willingness to confront the limitations of the field, may indeed be the “silver lining” of the Mayfield case. See Jennifer Mnookin, Op-Ed., The Achilles’ Heel of Fingerprints, WASH. POST, May 29, 2004, at A27.


The Spanish authorities were less convinced, and after several weeks, located another suspect, an Algerian named Ouhnane Daoud, in a different database, who, they claimed, was the actual source of the print. Eventually, the FBI concurred. The FBI was deeply embarrassed, Mayfield was released from custody, and eventually received compensation of $2 million.\footnote{Eric Lichtblau, U.S. Will Pay $2 Million to Lawyer Wrongly Jailed, N.Y. TIMES, Nov. 30, 2006, at A18.}

What happened? I will mention just two of the most important causes of the error. First, one portion of one of Brandon Mayfield’s prints really does bear a striking resemblance to one portion of one of Ouhnane Daoud’s fingers.\footnote{OIG REPORT, supra note 52 at 6-7.} There is no doubt that portions of the two prints are extremely similar, and the resemblance between Mayfield’s finger and the portion of the image most clearly visible on the latent recovered from Madrid was, as it happens, particularly strong. How often are we likely to see such a high degree of resemblance in prints from different sources? No one really knows. The Inspector General’s report, an independent investigation conducted in the aftermath of the scandal, insists that this degree of similarity is extraordinarily rare.\footnote{Id. at 7.} Perhaps so, but the truth is that we do not actually know how common or rare that degree of apparent similarity may be. It is clear that the growing size of the databases used for fingerprint analysis increase the risks of misidentifications like this one.\footnote{Dror & Mnookin, supra note 17, at 55-56.} Latent fingerprint examiners, at present, do make regular use of AFIS systems, computerized databases to generate a set of possible matching prints—possible “cold hits” based on a fingerprint match. But the computer algorithms are far from perfect, and thus the computer search process alone cannot determine whether any of the possible prints actually match. Only the examiner, using ACE-V to compare each AFIS suggestion to the latent print, can make that determination. As the databases grow, so grows the possibility of highly similar near-misses like Brandon Mayfield’s—fingerprints so similar that they might fool even crack fingerprint experts.\footnote{Id.}
In addition to the unexpected degree of similarity between the prints from different sources, it appears that cognitive bias also played a role in the debacle. Immigration lawyer Mayfield was a Muslim; he had converted to Islam some time earlier. He had also once represented a known terrorist in a child custody dispute. While it appears that the FBI investigators did not know these facts about Mayfield when they first determined that his print matched the one found on the detonation materials in Madrid, their subsequent awareness of this information made them more reluctant to reopen the issue or contemplate the possibility that they had made an error. More generally, even apart from this contextual information, it seems that once the first FBI examiner declared the prints to match, the verifying examiners expected to find a match. It is no great surprise, then, that they found precisely what they expected to find, likely the result of a mixture of peer pressure and expectation bias.

This problem of biasing information goes well beyond the Mayfield debacle. Forensic experts frequently have access to information about a case that goes beyond whatever information is actually necessary for their forensic testing. They may be told by detectives or investigators about other powerful evidence linking the suspect to the crime. They may be told details about the suspect that bear no relation to the pattern identification evidence itself—that he is a known gang member, or that she has prior convictions, or that he has confessed, or that this match is critical because it is the only strong evidence in the case. No information of this kind bears in any way on the actual forensic science inquiry, and risks creating an unconscious biasing effect on the examiner. Indeed, in most scientific fields, there is a careful and often formalized effort to shield researchers from this kind of contextual information. It’s too dangerous. We human beings have a cognitive tendency to see what we expect to see. Think of the way that medical researchers make use, whenever possible, of carefully controlled studies to ensure that not even the treating physicians know who is receiving the medication under investigation and who may be getting the placebo. These

---

59 OIG REPORT, supra note 52.
protections exist to protect physicians from unconscious bias that might influence their interpretation of the effects of either the medicine or the placebo. Currently, in the forensic sciences, there are generally no such procedures to protect examiners from extraneous information that may have an unconscious influence on their findings.62

To be sure, some information, though potentially biasing, may nonetheless be necessary for conducting the forensic test. A fingerprint examiner, for example, will likely need to know what surface a print came from, notwithstanding that the information may provide context clues about the crime itself. The point is not that examiners should lack all access to non-forensic information relating to the case. Rather, to the maximum extent practicable, they should only be given the case-related information that is actually relevant and helpful to their forensic inquiry. Dan Krane and others have coined the name “sequential unmasking” as a label for this approach, in order to emphasize that forensic analysts ought to learn only that information that they actually need, and only when they actually need it. All information, in other words, should be unmasked—that is to say revealed—to the examiner in sequence, and only when it is necessary.63 The examiner should have access to all the information necessary to do his or her analysis effectively—no more, and no less.

The concern about the danger and power of biasing information is not simply theoretical. In a clever experiment, cognitive psychologist Itiel Dror used the Mayfield case to show the possibility of contextual bias effects on fingerprint examiners’ interpretations. A small handful of fingerprint examiners were each given a pair of prints, a latent print and a potential source print, and told that they were the prints from the Mayfield case. Each examiner was asked to evaluate whether or not the prints matched, using only the information contained in the print.64

62 See generally id; see also, e.g., Michael J. Saks et al., Context Effects in Forensic Science, 43 SCI. & JUST. 77 (2003).
In fact, however, unbeknownst to the examiners, the prints were not the Mayfield prints. Each examiner was actually given a set of prints that he or she personally had previously testified in court were a 100% certain, positive, error-free individualization. But now, when provided with this biasing contextual information suggesting that the prints were those involved in the Mayfield scandal, 60% of the examiners (three of the five examiners tested) reached the opposite conclusion, determining that the two prints in front of them did not in fact match.65 A fourth examiner judged the prints to be inconclusive.66 Only one of the five examiners reached a conclusion consistent with his or her original judgment that the prints matched.67 To be sure, the Mayfield incident was a significant scandal, so the potential biasing effect of this context information was obviously quite extreme. Nonetheless, given some fingerprint experts’ insistence that their methodology is not vulnerable to unconscious bias or general human fallibilities, Dror’s findings generated a great deal of interest and a certain amount of both surprise and anxiety within the fingerprint community.68 The experiment was, in a sense, a possibility proof, showing that bias could indeed, at least in some circumstances, be significant enough to affect examiners’ conclusions. Follow up experiments by Dror and his collaborators on a larger number of examiners and with less starkly biasing information still revealed the potentially biasing effect of contextual information on analysts’ judgments.69

From one perspective, these findings are quite unsurprising. Research across a variety of other fields shows that we are all potentially biased by context and expectation.70 Given that, why should we be in the least surprised that forensic science is no different than other cognitive enterprises?

65 Id.
66 Id.
67 Id.
70 See generally Riser et al., supra note 62, and the many sources cited therein.
The recognition that fingerprint examiners are potentially subject to bias does not mean that they are behaving unprofessionally, or being careless, or not trying hard enough. Cognitive biases are an inherent danger of our cognitive architecture. That forensic scientists are not immune to them is hardly a surprise—except, perhaps, to those forensic scientists who were committed to a conception of their infallibility.

Dror's studies do make absolutely clear that bias is not simply a theoretical concern but a practical one as well. Moreover, they reveal that at least in some circumstances, biases may be acute enough to affect forensic examiners' outcome judgments. These studies therefore suggest both the theoretical and practical importance of taking concrete steps to limit examiners' access to biasing information. To be sure, these studies are preliminary, and involve only a relative handful of examiners. Nonetheless, these studies, coupled with all that is already known about bias in other cognitive domains, strongly suggest that this area should receive significant further study, so that we can better understand how and when bias poses a danger, and how often bias may be strong enough to affect an examiners' conclusions. There is also a significant need to identify the mechanisms that could usefully reduce those biasing effects that cannot be eliminated.

Furthermore, as law professor Michael Risinger has pointed out, the limited forensic science research on biasing effects to date ought not to be taken as an excuse for inaction. We know enough right now that we ought not to require further research before taking all reasonable action to reduce bias effects. The existing forensic-oriented studies, albeit limited in number and preliminary, coupled with the far more substantial research in other domains revealing human beings' cognitive vulnerability to bias, should reverse the burden of proof: unless the forensic science community can establish that it does not need blinding protocols, masking procedures, or other mechanisms that would reduce or eliminate bias, we should assume that bias-reduction mechanisms are indeed already warranted by our current degree of knowledge.

---

71 Id.
73 See generally id.
II. THE NATIONAL ACADEMY REPORT ON THE STATE OF FORENSIC SCIENCE

For those who closely follow the debates in forensic science, much of what I have described up to this point will have been quite familiar. But for those who began reading this Article believing in the television version, so to speak, the significantly more complicated reality may have come as something of a surprise. I hope I have, in this whirlwind tour, left those of you who walked with great confidence in the reliability of fingerprints at least a bit unsettled. Unfortunately, for the most part, the same cannot be said for the courts.

Beginning around a decade ago, enterprising attorneys began bringing admissibility challenges to fingerprinting, drawing to the courts’ attention the kinds of problems and weaknesses I have just described. But nearly all of these challenges failed. One lone court did actually exclude fingerprint evidence in 2007, calling it “a subjective, untested, unverifiable identification procedure that purports to be infallible.” But although that case parroted what had practically become conventional wisdom in certain academic circles, as a legal decision it was a voice in the wilderness. Every other trial judge who has considered the admissibility of latent fingerprint identification since Daubert has found it to meet the applicable standard for the admissibility of expert evidence. It has been, for the most part, too hard for judges to contemplate excluding a form of evidence that has been routinely used for nearly a hundred years. For those academics engaged in constructive criticism of the forensic sciences, reading these judicial opinions has often felt like walking in a house of mirrors. As I will describe in more detail below, evidence of the problems facing forensic science would often be ignored, distorted, or recast, in order to help courts avoid

confronting the insufficient research basis supporting these forms of evidence.\footnote{See infra Part III.}

Meanwhile, late in 2005, partly at the request of the leadership of the forensic science community itself, Congress commissioned the well-regarded, independent, and non-partisan National Academy of Sciences to research and write a report on the needs of the forensic science community. In February of 2009, the long-anticipated report was issued.\footnote{NAS Report, supra note 9.}

This report was written by an interdisciplinary panel of distinguished scholars and practitioners, who conducted their own investigation into the state of the research, and also heard numerous days of testimony from a substantial number of leading forensic science professionals, researchers, and others knowledgeable about the state of the forensic scientists.\footnote{In the interests of full disclosure, I should note that I provided oral commentary to the NAS panel during their process.} This panel included scientists from a variety of fields, several forensic professionals, and some with legal experience.\footnote{The make-up of the committee has been criticized by the forensic science community for having insufficient practitioners. It is interesting to note that the committee in fact had several forensic practitioners, while it had not a single member who had already published critical work concerning the adequacy of the research basis of forensic science. For more information on the committee and its make-up, see generally D. Michael Risinger, The NAS/NRC Report on Forensic Science: A Path Forward Fraught with Pitfalls, UTAH L. REV. (forthcoming).} The panel also included one law professor: Brooklyn Law School’s own Margaret Berger, in whose honor this Festschrift volume of the Law Review has been written. In essence, the 319-page report substantially confirms the views of the academic critics about the inadequacy of the research basis to support many of the claims routinely made by forensic scientists.\footnote{NAS Report, supra note 9.}

For example, the report finds that there is not an adequate basis for claims of individualization. The report also finds “a notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods.”\footnote{Id. at 5-6.} The report claims that research on proficiency, performance, and the role of bias and observer effects is “sorely needed.”\footnote{Id.} All in all, “[t]he present situation . . . is seriously wanting, both because of the limitations of the judicial system.
and because of the many problems faced by the forensic science community.\textsuperscript{82}

The report's boldest and perhaps most important recommendation is for the creation of a new independent federal agency to regulate, supervise, and improve the forensic sciences. This agency, dubbed the National Institute of Forensic Science (NIFS), would be responsible for funding research to improve forensic sciences; it would also be responsible for establishing and developing best practices, and, more generally, supporting and overseeing the forensic science infrastructure.\textsuperscript{83} One academic has quipped that the NIFS would be “a mixture of the SEC and the NSF” for forensic science, rather an apt characterization.\textsuperscript{84} The report explains in detail that no other existing agency has, in the committee's view, the ability effectively to provide all of what forensic science needs in terms of both research management and regulatory oversight.\textsuperscript{85} No existing agency—neither NIST, the NIJ, nor anyone else—has, according to the report, the necessary expertise, resources, and appropriate political culture to permit it to perform this array of functions credibly and successfully.\textsuperscript{86}

The report makes a number of other significant recommendations. It calls in strong terms for additional research to establish the validity and reliability of forensic sciences, as well as research to examine the extent of biases and observer effects.\textsuperscript{87} It calls for mandatory laboratory accreditation and mandatory individual certification of forensic scientists (right now both are entirely optional).\textsuperscript{88} Significantly, it calls for the use of incentive funding to motivate states to make their crime laboratories independent from law enforcement and prosecutors.\textsuperscript{89}

The report has received a good deal of attention both from within the forensic science community and from outsiders. As of now, several hearings on Capitol Hill have been held to

\hspace{1cm}\textsuperscript{82} Id. at 5-9.
\textsuperscript{83} Id. at 5-14.
\textsuperscript{84} Roger Koppl, Professor of Econ. & Fin., Farleigh Dickinson Univ., Remark at a Forensic Science Conference at Arizona State University Sandra Day O'Connor College of Law (2009).
\textsuperscript{85} NAS Report, supra note 9, at 5-24.
\textsuperscript{86} Id.
\textsuperscript{87} Id. at 5-18.
\textsuperscript{88} Id. at 5-19.
\textsuperscript{89} Id. at 5-17.
consider its recommendations.\textsuperscript{90} President Obama has created a forensic science task force to consider how to proceed.\textsuperscript{91} But the unfortunate reality is that now, roughly a year after the report was issued, almost no one believes that NIFS is going to be on the horizon any time soon. A new federal agency seems to be neither fiscally nor politically viable.

Overall, the NAS report is an impressive achievement, and both its criticisms and suggestions are, in my view, generally on the mark. However, I do have one relatively significant quarrel with the report, which relates to its treatment of the legal aspects of forensic science and admissibility. The report offers a thorough and trenchant critique of how the courts have thus far handled forensic science. It describes how the courts have substituted long use for an actual focus on proven validity.\textsuperscript{92} It points out the ways in which judges have been “utterly ineffective” at honestly assessing the research basis of the pattern identification sciences.\textsuperscript{93} I agree completely. The report further contends that the judiciary, particularly given judges’ lack of training in science, the case-by-case nature of the adjudicatory system, and the limits of appellate review, cannot be expected to solve this problem on its own. “Judicial review, by itself, will not cure the infirmities of the forensic science community.”\textsuperscript{94} Again, I wholeheartedly agree.

But then, at this key point, the NAS report decides to punt. After offering this significant critique of the judiciary’s actions, the report is distressingly silent about what the judiciary ought to do next. If judicial review by itself will not and cannot solve our problems in the forensic science arena, does it nonetheless have some role to play? While we await NIFS, or any other significant federal initiative vis-à-vis forensic science; while we await the necessary research that we hope will eventually be done; while we await greater regulation and the possibility of mandatory accreditation, what should

\begin{itemize}
  \item \textsuperscript{90} Strengthening Forensic Science in the United States: Hearing Before the S. Comm. on the Judiciary, 111th Cong. (2009), webcast available at http://judiciary.senate.gov/hearings/hearing.cfm?id=4038. The Congressional Subcommittee on Technology and Innovation also held a hearing on March 10, 2009 and the Judiciary Committee also held a hearing on May 13, 2009.
  \item \textsuperscript{91} The National Science and Technology Council in the Office of the President of the United States created a Subcommittee on Forensic Science. For details, see http://projects.nfstc.org/trace/2009/presentations/7-melson-stolorow-nas.pdf.
  \item \textsuperscript{92} NAS Report, supra note 9, at 5-9.
  \item \textsuperscript{93} Id.
  \item \textsuperscript{94} Id. at 12.
\end{itemize}
judges do? Sit idly by, and continue permitting pattern identification evidence as they always have? While providing a good deal of persuasive authority to any judge who wishes to scrutinize forensic science with more care, the NAS report does not offer any specific guidance whatsoever for courts wrestling with admissibility determinations under *Daubert*.

Understandably, the committee may have thought that sweeping statements about admissibility were unwarranted, and more generally that telling judges how to behave would be overstepping its mandate. I fully recognize that the NAS report was directed at Congress and was not primarily intended for the courts. But the committee might have done more. For example, while the report says there is no scientific basis for claims of individualization, it does not go one step further and explicitly say that such testimony ought therefore to be deemed objectionable. The report says that a zero error rate is not scientifically plausible, but it does not explicitly say that testimony asserting such a rate should therefore be viewed with skepticism.

To be sure, these are easy inferences to draw from what the report *does* say. If *Daubert* says that expert evidence needs to be established as reliable and valid, and a court accepts the NAS report’s conclusion that a zero error rate is not scientifically plausible, then it might be a simple matter of syllogistic logic to conclude that any claim of a zero error rate does not pass *Daubert*. And there is no doubt that the report is already being cited by defense attorneys in the latest round of *Daubert* motions that are, inevitably (and, in my view, quite appropriately) being spurred by the report’s contents and conclusions.

But there is, in the report, a certain sense of resignation, perhaps even fatalism, about the courts. There is a tone that suggests that the committee may have thought that the judiciary has done such a poor job of gatekeeping in this area that it is hard to imagine the possibility of change. To have faith in the courts’ capacity to make an about-face in this domain is, perhaps, the mark of a naïve optimist, or maybe even a chump.

While I agree with the report that courts alone cannot and will not prompt the necessary reforms to forensic science,

---

95 Id.
96 See generally id.
it also seems to me that courts absolutely do need to be part of the solution. For courts to continue to treat forensic science evidence in the same manner that many of them did prior to the NAS report would be worse than cowardly. The NAS report strongly suggests that the concerns expressed in recent years by a number of academic critics were neither overstated nor illegitimate. The academic critics were not crackpots, nor were they Chicken Littles worrying about something that no one else could see because it wasn’t really there. There truly is an insufficient basis in research for many of the strong claims that forensic scientists have been making for years. The courts should squarely confront that fact and decide in a thoughtful way what consequences it creates for these forms of evidence; what effects it has on admissibility; and whether it means that the evidence should, at a minimum, be limited, and if so, how.

Admittedly, the question of whether fingerprint evidence should pass Daubert or the equivalent test under state law is a difficult one. Fingerprint evidence has been used in court for roughly one hundred years. There is no doubt that the pattern variation among human friction ridges is indeed enormous, and that fingerprints, whether or not they are truly unique, certainly have extremely significant discriminatory power. Fingerprint evidence is quite obviously probative. Moreover, people trust it. Not just “people”—not just the proverbial man in the street, not just forensic scientists themselves, not just prosecutors. Even those of us—like me—who are sometimes viewed as critics of fingerprint evidence acknowledge its probative power. For example, if you asked me which piece of evidence I would have more confidence in, an eyewitness identification by a crime victim of the perpetrator, a stranger who was viewed for a short period of time by the victim during a stressful crime; or a latent fingerprint identification of an individual made from several high-quality, clear latent prints found in a location and in circumstances strongly suggesting that whoever left the print was connected to the crime, my honest answer would be the fingerprint identification. All of these facts make it extremely difficult for a court to seriously and deeply consider the possibility of excluding, or even limiting, this form of evidence.

97 Indeed, even Simon Cole and Michael Saks, two of the staunchest so-called critics, both recognize that fingerprint evidence has a great deal of power. See, e.g., Cole, Grandfathering Evidence, supra note 1, at 1193-94; Saks, Merlin and Solomon, supra note 1, at 1106.
And yet, in a way these facts about the power of fingerprint evidence are beside the point, or at least they do not adequately and fully answer the question of fingerprint evidence’s legitimate admissibility in court. The problem with fingerprint evidence is not that it completely lacks probative power, but rather that research on the domain has not yet established the appropriate limits to its probative power, or shown how that value varies depending on its quality or its quantity of information. It is as if I wanted to sell you a valuable jewel and the question was how much it should cost. We may both agree that the jewel is indeed valuable, but the right question is not whether it has value, but just how valuable it is, and just how we can know its appropriate value. If I refused to research that question, or even to permit any research on the specific characteristics of the jewel that would help assess its precise value, and I further refused to research the relevant market information to help determine its appropriate price, I would have an extremely hard time selling the jewel, and legitimately so. If all I said was, “Of course it’s valuable. We both know it’s simply the most valuable jewel in the world. I’m not going to investigate the question of its precise value, but we both know it’s virtually perfect. Let’s therefore just agree that it’s worth an extraordinary amount and price it accordingly,” we would both expect you to walk away from the potential sale.

Fingerprint evidence, like that jewel, is obviously valuable. But like that jewel, we should be wary of ‘buying’ it as legal evidence in court until we have a better, research-based understanding of precisely how valuable it is, and whether there are instances in which we might be assuming it to have a significantly higher value than it really does. Without careful proficiency testing of examiners, without information about what the significance of any given ‘match’ really is, without error rate information about the frequency and circumstances of mistakes, without understanding which fingerprint identifications are easy and which ones are more difficult and hence more likely to be error-prone, we should be cautious buyers indeed. Quite possibly, the better strategy might be simply to forego the purchase until more of the necessary information is available. Still, recognizing this somewhat awkward state of affairs—that we have a form of evidence that obviously often has probative power, but also has a strikingly inadequate research basis—helps, I think, to make sense of the courts’ reluctance to exclude, or in some
circumstances, even honestly to engage with the arguments made by defense counsel.

III. HOW HAVE THE COURTS ANALYZED THE ADMISSIBILITY OF FINGERPRINTS?

In the remaining portion of this article, I will explore both what courts have actually done when confronted with challenges to the admissibility of pattern identification evidence, as well as what they ought to do. First, I will focus on what they have done, and then turn to what they ought to do, but it may be valuable to preview my suggestions about what they ought to do before turning to look at how courts have actually responded to the numerous pattern identification admissibility challenges that have arisen.

My bottom line is straightforward: Forensic science experts should not continue to be given free rein to testify in the manner they have typically done up until now. Judges need to develop a variety of thoughtful approaches—a toolkit of sorts—with which they can assess admissibility, and this toolkit should absolutely include outright exclusion in some circumstances. What judges ought to do may well not be the same across the board, even within the same field—as Kumho Tire v. Carmichael\textsuperscript{52} indicates, and as Michael Risinger has usefully emphasized,\textsuperscript{53} the court’s responsibility when assessing the admissibility of expert evidence is to focus on “the task at hand”—which means looking closely at the specific nature of the claims being made in the particular circumstance. This means that we may not be able to achieve—or may we necessarily want—a one-size-fits-all admissibility approach to any given form of pattern identification evidence. Admissibility determinations, and the scope of permissible testimony, may depend on the details—is it a single partial latent from AFIS, or are there nine extremely clear latents corresponding to three different fingers of the defendant? What is the quality of the print or prints in question? What specific claims is the expert trying to make, and how absolute are the conclusions presented? These questions, and others like them, ought to inform judges’ analyses. All pattern identifications are not

\textsuperscript{52} 526 U.S. 137 (1999).

created equal, and a blanket approach to admissibility, pro or con, is not likely to be warranted.

Nonetheless, at a bare minimum, given currently available knowledge, courts ought not to permit evidence of either individualization or a zero error rate. Even the forensic science community is beginning—slowly and in the case of individualization somewhat falteringly—to recognize that these claims are perhaps better let go. But I would go further. In many cases involving pattern identification, courts have, as of now, only two legitimate choices if they are to take Daubert seriously: either (1) limiting the evidence by restricting it to description of similarities and differences, rather than offering opinions; or (2) outright exclusion. One alternative is to greatly restrict the expert’s testimony, limiting it to description of similarities and differences without any evidence providing an interpretation of these similarities and differences. This option, though superficially appealing, raises some thorny admissibility concerns which are hard to avoid in a principled way, as I will describe below. Nonetheless, it seems acceptable as an imperfect, interim solution for courts who simply find it too hard to exclude evidence that likely does have significant probative value. The second approach, and one that deserves far more serious consideration than most courts have been prepared to give it, is outright exclusion.

While that may sound like an extreme reaction—cutting off the patient’s arm when perhaps it could have been put in a cast and saved—it is important to note, as I will argue below, that with a modicum of effort on the part of researchers and forthright cooperation from the forensic science community, exclusion would be quite short-lived. As I will argue in Part IV,

---

110 On February 19, the President of the IAI, Bob Garrett, issued a two-page statement to its membership stating, “Although the IAI does not, at this time, endorse the use of probabilistic models when stating conclusions of identification, members are advised to avoid stating their conclusions in absolute terms when dealing with population issues.” Though the wording is perhaps slightly opaque, this is, in essence, a pulling back from an absolute claim of individualization. See http://www.theiai.org/current_affairs/nas_memo_20090219.pdf (last visited April 17, 2010). However, a month later, in written comments prepared for the Senate Judiciary Committee, the IAI backed off from this claim, and in a significantly longer document (7 pages of remarks), pointedly said nothing about individualization, whether conclusions can be absolute, or how to think about population issues. See http://www.theiai.org/current_affairs/nas_response_leahy_20090318.pdf (last visited April 17, 2010). For another example of a staunch defender of the reliability of pattern identification science suggesting that, for instrumental rather than epistemic purposes, claims of individualization should be dropped, see John Collins, Stochastics—The Real Science Behind Forensic Identifications (2009), Crime Lab Report, available at http://www.crimelabreport.com/library/monthly_report/11-2009.htm.
courts’ primary focus under *Daubert* should be on the question of whether the expert’s claims have been subject to adequate testing—whether there is evidence that supports the claim that the expert can actually do what she says she can do. Complete knowledge of the cognitive practices of fingerprint evidence; a validated statistical model of ridge characteristics and frequencies; even objective standards for determining whether or not prints match—all of these forms of knowledge would be valuable, beneficial, and are absolutely worth pursuing, but they need not exist in order for latent fingerprint evidence to be legitimately admissible under *Daubert*. What courts really ought to consider requiring, in many cases, as a *minimum prerequisite to admissibility* is simply much better error rate information about examiners’ abilities in practice. And producing this information is eminently achievable with concerted focus and effort.

I turn, now, to what courts have actually done when confronted with admissibility challenges to fingerprint evidence and other kinds of pattern identification evidence. Up to the present, courts wrestling with these admissibility challenges have offered several categories of arguments and made certain repeated intellectual moves. In this Part, I will describe three dominant analytic approaches that courts have taken: (1) the ostrich maneuver, (2) the ACE-V conclusion (where ACE-V stands, in this case, for “Admissible—Considering Everything, it’s Valid (enough));” and (3) the Solomonic compromise.

A. The Ostrich Maneuver: Problem? What Problem?

The first approach taken by some courts when confronting challenges to the admissibility of pattern identification evidence is what I call the ostrich maneuver, because these courts appear to be trying desperately to keep their heads in the sand. Certain judges have, to a sometimes remarkable extent, averted their eyes to the quite legitimate concerns about the research basis supporting the conclusions offered in the pattern identification sciences. These judges have more or less asked, “Problem? What problem?” An archetypal example of the ostrich maneuver occurred in *United States v. Havward*. The judge’s view of the issue was captured early in

---

111 117 F. Supp. 2d 848 (S.D. Ind. 2000). Though *Havward* is an especially dramatic “ostrich” opinion, it is certainly not the only one. See, e.g., *United States v.*
his written opinion finding that fingerprinting passed Daubert, when he wrote, “The court’s decision may strike some as comparable to a breathless announcement that the sky is blue and the sun rose in the east yesterday.”

The judge went on to explain why, in his opinion, fingerprint evidence posed virtually no genuine difficulty under Daubert. He concluded, for example, that fingerprint evidence could be relied upon because it had indeed been tested (testing being one of the most important factors under Daubert) “in adversarial proceedings with the highest possible stakes—liberty and sometimes life.” Note the judge’s slippage here—Daubert envisions scientific testing, not courtroom testing. In fact, the entire purpose of Daubert—a heightened reliability screen for expert evidence—derives from the idea that the crucible of the courtroom is, by itself, an insufficient check on validity and reliability for scientific and expert evidence. If testing through adversarial proceedings were enough, then a separate judicial inquiry into reliability would be entirely superfluous. We wouldn’t need Daubert at all, because the crucible of the courtroom would suffice. Notwithstanding the


Id.

To be sure, one could make the argument that we do not actually need a special judicial check on validity and reliability for expert evidence. One could possibly defend this claim on a variety of grounds, ranging from (1) an argument that adversarial testing through the presentation of contrary evidence and cross-examination are in fact an adequate method for evaluating expert evidence; to (2) that there may be little evidence that judges’ evaluation of the legitimacy of expert evidence is epistemically superior to that of juries, in which case using them as a ‘reliability’ screen for evidence might be epistemically ineffective and, therefore, potentially usurping the power of the jury for no legitimate purpose; to (3) that we do not engage in strong validity checks on other kinds of potentially unreliable evidence and there is an insufficient institutional or epistemic justification for treating expert knowledge differently, etc. But none of these arguments—whether or not they have any intellectual merit—provide a legitimate basis for a lower court within our legal hierarchy to argue that ‘adversarial’ testing constitutes testing of the sort envisioned by Daubert. A judge could make a more subtle argument that adversarial testing is adequate for “shaky but admissible” evidence. Daubert says, and indeed the judge in Havvard quoted, later in his opinion: “Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.” Havvard, 117 F. Supp. 2d at 854 (internal quotation marks omitted) (quoting Daubert, 509 U.S. at 596). But while the power of adversarial testing might be an argument for permitting borderline evidence (and indeed, is a contributing argument to the ACE-V approach to these questions discussed below, infra notes 114-128 and accompanying text), adversarial testing should not be understood as the kind of testing referred to and expected by Daubert.
sleight of hand at work in Havvard with respect to the idea of testing, adversarial testing simply cannot legitimately be seen as fulfilling Daubert’s idea of testing.

The judge engaged in a similar, though less egregious, sleight of hand with respect to the peer review and publication factors of Daubert. He granted that the “publication” factor was an awkward fit for fingerprint evidence because it had been developed “for forensic purposes,” but found that adversarial testing offered an adequate substitute. He also found that there was plenty of peer review, both because of the practice of having one examiner verify the conclusions of another, and also because any examiner could review the conclusions of any peer by taking her own look at the prints in question.

Is the verification stage of ACE-V akin to peer review? It is true enough that having verification as a standard practice does provide a certain degree of routine peer examination for each declared fingerprint match. But peer review of, for example, manuscripts for publication, usually, though not always, makes the reviewers blind to authorship. More often than not, peer reviewers do not know whose manuscripts they are reading so that they are not potentially biased by the authors’ credentials or experience. Moreover, peer review in the context of research assessments permits a kind of semi-public scrutiny and dialogue regarding the merits of the researchers’ methodology and approaches, precisely because it typically involves scrutiny by other experts from outside the original author’s workplace or close circle of collaborators. Having your immediate colleagues take a look at an article is not what journals—or Daubert—means by peer review. Moreover, peer review usually also requires written commentary and reasoned explanation.

The verification phase of fingerprint evidence thus fails to map precisely onto scientific peer review in several important ways. To be sure, if verification were conducted as a “blind” review, i.e., the verifying examiner did not know whose identification she was verifying or even that she was conducting a verification rather than an initial analysis, it

---

105 Havward, 117 F. Supp. 2d at 854.
106 Id.
would be a closer, though still not exact, fit. But even blind peer review would be more of a check on consistency—reliability in the scientific, rather than legal sense—rather than validity or accuracy. That two examiners reached the same conclusion shows no more and no less than that they reached the same conclusion. If blind verifiers consistently reach the same conclusion as the original analyst, it shows that the method generates consistent results, but it does not necessarily show that it achieves accurate results, though in some instances we may be willing to infer accuracy from consistency.\footnote{Accuracy and consistency are not necessarily correlated—I could have a miscalibrated instrument that gives me the same result every time and yet is always wrong. It is equally possible that there could be a method that generates consistent results even when practiced by different individuals, but these consistent results were not accurate. Think for example, of a poorly designed psychological test for depression. Different testgivers might use the test on patients, get consistent diagnoses, and yet be wrong, because the test wasn’t well designed in the first place. Ironically, in the case of latent fingerprint identification, the absence of a consistent method (given that, as discussed above, at infra notes 22-23 and accompanying text, ACE-V is less a method than an outline of procedural steps for careful looking) might legitimate an inference of some degree of validity from generally consistent results through blind peer review. If examiners have extremely consistent results notwithstanding the fact that they use an array of imprecisely defined and not-quite-identical approaches, including choosing different minutiae, different standards for determining the quantum of information required to call an identification, and varied practices for conducting a comparison, we might be prepared to think their consistent results provided some warrant for inferring validity. It is as if instead of conducting the same psychological test multiple times, we conducted several different psychological tests and all of them resulted in the same diagnosis; these multiple methods' consistent results would be cumulative, and thus could increase our confidence in the correctness of the original diagnosis. The irony is that this potential evidentiary power of verification results from the lack of formalized or consistent standards; to whatever extent examiners are doing the same thing as one another, verification can establish only consistency, or a check on methodological error, but not validity, because the verification evidence would be merely redundant, rather than cumulative. On cumulative and redundant evidence, see generally DAVID A. SCHUM, THE EVIDENTIAL FOUNDATIONS OF PROBABILISTIC REASONING (2001). For an analogous argument in the context of human/technology partnerships in fingerprint evidence, see Dror & Mnookin, supra note 17, at 60-65.}

Clearly, then, verification cannot simply be equated with academic peer review. However, I grant that equating verification with peer review is not quite as distressing, conceptually, as equating adversarial testing with scientific testing, especially given that Daubert intended its criteria as suggestions rather than hard requirements. These suggestions may be appropriately modified to apply to the particular circumstances of the practices at issue and the “task at hand” and thus a loose analogy between peer

\footnote{\textit{Kumho Tire}, 526 U.S. 137, the third case in the so-called \textit{Daubert} trilogy, emphasizes the importance of focusing on the “task at hand.” For discussion of the}
review and verification might not be out of place. But given that there is absolutely no publicly-available data about how often peer review of the sort engaged in by fingerprint examiners actually catches errors, or how well it functions in operation, the courts’ easy willingness to presume its utility seems to be overreaching. For example, if non-blind verifiers virtually never disagree with the original analyst, than the power of this “review” as a check may in fact be quite limited, for it might be that whatever errors do occur are, as in the Mayfield case, likely to be repeated by subsequent analysts as well.  

But while the analysis in Havward is generally unpersuasive, it is the opinion’s conclusion that is most ostrich-like of all. The court wrote, “In sum, despite the absence of a single quantifiable standard for measuring the sufficiency of any latent print for purposes of identification, the court is satisfied that latent print identification easily satisfies the standards of reliability in Daubert and Kumho Tire. In fact, after going through this analysis, the court believes that latent print identification is the very archetype of reliable expert testimony under those standards.” Reasonable people can disagree about whether fingerprint evidence ought to be admissible under Daubert. But to argue that it is the “very archetype of reliable expert testimony under those standards” strains all credibility. Only by putting one’s head in the sand could one possibly conclude that latent fingerprint evidence—which has been tested in the adversarial crucible but not scientifically, lacks meaningful error rate information, and operates without statistical foundation or any validated, objective criteria for determining a match—is the archetype of reliable evidence under Daubert.

---

110 However, if verifiers disagree with the original analyst more than occasionally, that provides indirect evidence that the error rate for an individual examination is not, in fact, negligible.

111 Havward, 117 F. Supp. 2d at 854 (emphasis added).

112 Id.

113 Another ostrich maneuver, in some ways even more extreme than the superficial and poorly-reasoned analysis in Havward, occurs when courts decide that pattern identification evidence is so clearly reliable that they do not even permit the defense to hold a preliminary hearing on the issue. See, e.g., United States v. Joseph, 2001 WL 515213, *1 (E.D. La. May 14, 2001), for one example. Of course, this issue is complicated by the question of what the defense actually proffers. For example, if a defense attorney asks for a hearing but has done no substantial preparation and fails to make a credible argument, a judge would appropriately deny the request. That
B. The Judiciary’s Own Version of ACE-V

The second approach is what I, slightly tongue-in-cheek, call the judiciary’s version of ACE-V. In this case, the acronym stands for “Admissible—Considering Everything, it’s Valid (enough).” This category of cases includes those courts who attempt, with at least a modicum of seriousness, to engage Daubert and Kumho Tire’s requirements vis-à-vis pattern identification testimony, and who recognize that the research basis supporting the evidence is not what one would wish it to be. Nonetheless, these courts, though squirming a bit and acknowledging some of the legitimate concerns regarding the research basis for this evidence, find that, on balance, the evidence still warrants admission in its traditional form, though without fully explaining what justifies this conclusion.

United States v. Sullivan is an illustrative example of this approach. The court wrote:

The court shares the defendant’s skepticism that the ACE-V methodology enjoys a 0% error rate, making it effectively a perfect art. There is no evidence, however, that the ACE-V methodology as performed by the FBI suffers from any significant error rate. The FBI examiners have demonstrated impressive accuracy on certification-related examinations, and Younce testified that an examiner who made a false identification would be finished as an examiner due to the difficulty in rehabilitating him or her as a witness. While the defendant is correct that the party submitting the evidence has the burden of establishing its reliability under Daubert, the defendant has failed to submit any evidence to dispute the plaintiff’s evidence of a minimal error rate. Consequently, while the court rejects the plaintiff’s claim of a 0% error rate, it finds that the

appears to be what occurred, for example, in United States v. Pena, in which the trial court reasoned, “the case law is overwhelmingly in favor of admitting fingerprint experts under virtually any circumstance.” 586 F.3d 105, 110 (1st Cir. 2009) (internal quotation marks omitted). “Consequently, the court reasoned, the only way it would have considered excluding the testimony or giving a limiting instruction ‘is if there had been data, real evidence presented about the limitations of fingerprinting.’” Id. “Instead, as the court acknowledged, Pena’s motion to exclude relied on ‘one article from the Fordham Law Review, and that’s not enough to carry the weight of the exclusion motion.’” Id. It is quite understandable that a court is not going to waste its time with a preliminary hearing when the defense is so ill-prepared. But the appellate case law in the case suggests that it is never an abuse of discretion to admit fingerprint evidence without a preliminary hearing, and this is a distressing—and ostrich-like—conclusion. Id. Interestingly, just as this Article was going to press, Judge Nancy Gertner, whose opinion refusing a Daubert hearing was affirmed in Pena, subsequently issued a novel procedural order explicitly referencing the NAS report, to make clear her openness to hearing well-prepared challenges to pattern identification evidence. See Procedural Order, available at http://www.mad.uscourts.gov/boston/pdf/ProcOrder TraceEvidenceUPDATE.pdf (last visited April 17, 2010).
error rate is not sufficient to render ACE-V unreliable under Daubert.\textsuperscript{115}

The court squarely recognized that a zero percent error rate is not credible. The court further acknowledged that the proffering party had the burden of showing reliability, which ought to have meant that a lack of error rate information counted against the prosecution.\textsuperscript{116} Nonetheless, in the face of this acknowledged absence of information supporting any particular error rate, the court found that good performance on certification exams and a culture that deems any discovered identification errors so impermissible as to be career-ending, spoke sufficiently to the question of error rate as to render the ACE-V approach to fingerprint identification reliable under Daubert. To be sure, in the right circumstances, proficiency tests could indeed provide substantial and adequate information regarding error rates, but the current certification tests are neither challenging enough nor taken in circumstances sufficiently mirroring actual practice for this inference to be warranted.\textsuperscript{117} Similarly, the fact that the professional culture within the latent fingerprint community may view any false identification error as career-jeopardizing if discovered does not itself establish that errors do not occur. Indeed, one could imagine that a professional culture that deems certain errors career-ending might also be a culture that chooses not to look very hard for them.

Turning to the testing factor of Daubert, the judge explained:

The court further finds that, while the ACE-V methodology appears to be amenable to testing, such testing has not yet been performed. The court disagrees that testing that establishes the validity of the principles underlying ACE-V—that fingerprints are unique and permanent—can substitute for testing of the ACE-V methodology itself. That testing, however, is relevant as it provides a foundation for the ACE-V methodology. . . . But as the defendant points out, there is not a standard defining how many similarities must be found before a match is declared. Younce testified that there is no minimum number of “points” in common necessary to declare a match between a known and an unknown print. Indeed, Younce testified that such a requirement would be unscientific. While it is possible that this position is ultimately correct, it is not supported by the studies submitted by the plaintiff. Evidence that no two

\textsuperscript{115} Id. at 704.
\textsuperscript{116} Id.
\textsuperscript{117} See generally Koehler, Proficiency Tests, supra note 1; Mnookin, Confessions of a Fingerprinting Moderate, supra note 1, at 136-37.
fingerprints are the same—or that no two 21.7% of a print are the same—is not evidence that no two fingerprints can share a partial print in common. The court finds that this concern does not render fingerprint evidence unreliable for the purposes of Daubert. While the possibility that two fingers may have a fractional portion of a print in common may affect the probability estimates that two fingers may leave the same fractional print, that possibility goes to the weight of the evidence, not its admissibility.\textsuperscript{118}

Note that here, the court determined that testing \textit{could} be done and that it \textit{had not yet} been done. Furthermore, the court understood, quite correctly, that evidence supporting the uniqueness and permanence of fingerprints did not itself establish that fingerprint examiners could actually make certain identifications. The court further recognized the lack of objective standards for determining a match. But at this point, the court essentially punted. Without any serious effort at analysis or explanation, the court simply decided that these difficulties went to the weight of the evidence rather than its admissibility.\textsuperscript{119} Still, compared to ostrich judges, this court at least elected to name clearly some of the problems with the research basis supporting fingerprint identification.

Another one of the many examples of judicial ACE-V thinking is the appellate opinion in \textit{United States v. Mitchell}.\textsuperscript{120} The court acknowledged that if “directed, specific actual testing” were the requirement of Daubert, then fingerprint evidence would have significant problems.\textsuperscript{121} However, the court seemed to believe that “directed, specific, actual testing” was not in fact required; instead, the long and substantial history of the use of the technique could provide what the court considered to be a form of “implicit testing.”\textsuperscript{122} Moreover, the court emphasized that while it was indeed required to be a “gatekeeper” for scientific evidence, it was also “only a gatekeeper, and a gatekeeper alone does not protect the castle.”\textsuperscript{123} All in all, though the evidence supporting validity might be imperfect, the real question, in the court’s estimation, was whether the available evidence of validity was sufficient to pass the baton from the court to the adversary system.\textsuperscript{124} In other words, fingerprint evidence might not be Valid (with a

\textsuperscript{118} Sullivan, 246 F. Supp. 2d at 704.
\textsuperscript{119} Id.
\textsuperscript{120} 365 F.3d 215 (3d Cir. 2004).
\textsuperscript{121} Id.
\textsuperscript{122} Id.
\textsuperscript{123} Id.
\textsuperscript{124} Id.
capital V) but it was valid enough, in the court’s estimation, for cross-examination and counter-expertise to be an adequate corrective to whatever limitations might be present.125

Another, and perhaps the best-known, example of judicial ACE-V thinking is Judge Louis Pollak’s second opinion in United States v. Llera Plaza (“Llera Plaza II”).126 The first time he considered fingerprint evidence, the judge—an especially thoughtful jurist, and a former professor and dean at Yale Law School—made fingerprint history by ruling, in Llera Plaza I, that latent fingerprint evidence did not pass muster under Daubert. After a re-hearing on somewhat unusual procedural grounds, Judge Pollak, by his own admission, changed his mind. In his second opinion, Llera Plaza II, he continued to find that Daubert’s testing factor was not fulfilled. On error rates, he found that the extant proficiency tests gave little assistance because they were too easy to provide a
discriminating measure of competence. But he nonetheless found that the absence of any evidence showing a substantial error rate provided some evidence that the error rate was tolerably low.\textsuperscript{127} His bottom line was, more or less, that fingerprint evidence wasn’t exemplary, but it was good enough: “[T]o postpone present in-court utilization of this ‘bedrock forensic identifier’ pending such research would be to make the best the enemy of the good.”\textsuperscript{128} All things considered, it’s valid (enough).

C. Dividing the Baby: The Solomonic Compromise

The third approach, to which I will give the most attention because it represents the most thoughtful judicial approach taken to date to assessing the admissibility of forensic science evidence in court, is what I am calling the “Solomonic compromise.” This approach reflects judges’ efforts to split the difference between admissibility and exclusion in a way that superficially seems to make sense, but becomes, I will suggest, increasingly problematic upon careful reflection. This approach permits the expert to testify about similarities and differences in the patterns at issue, but prohibits or limits the expert from reaching expert conclusions about the meaning of those similarities. Llera Plaza I—the opinion Judge Pollak vacated after a preliminary hearing—took precisely this approach. His explanation in his original opinion is worth quoting in detail:

Since the court finds that ACE-V does not meet Daubert’s testing, peer review, and standards criteria, and that information as to ACE V’s rate of error is in limbo, the expected conclusion would be that the government should be precluded from presenting any fingerprint testimony. But that conclusion—apparently putting at naught a century of judicial acquiescence in fingerprint identification processes—would be unwarrantably heavy-handed. The Daubert difficulty with the ACE-V process is by no means total. The difficulty comes into play at the stage at which, as experienced fingerprint specialists Ashbaugh and Meagher themselves acknowledge, the ACE-V process becomes “subjective”—namely, the evaluation stage. By contrast, the antecedent analysis and comparison stages are, according to the testimony, “objective”: analysis of the rolled and latent prints and comparison of what the examiner has observed in

\textsuperscript{127} Id. at 566.
\textsuperscript{128} Id. at 572. For a thoughtful critical analysis of this opinion, see generally David H. Kaye, The Nonscience of Fingerprinting: United States v. Llera-Plaza, 21 QUINNNIPAC L. REV. 1073 (2003).
the two prints. Up to the evaluation stage, the ACE-V fingerprint examiner’s testimony is descriptive, not judgmental. Accordingly, this court will permit the government to present testimony by fingerprint examiners who, suitably qualified as “expert” examiners by virtue of training and experience, may (1) describe how the rolled and latent fingerprints at issue in this case were obtained, (2) identify and place before the jury the fingerprints and such magnifications thereof as may be required to show minute details, and (3) point out observed similarities (and differences) between any latent print and any rolled print the government contends are attributable to the same person. What such expert witnesses will not be permitted to do is to present “evaluation” testimony as to their “opinion” (Rule 702) that a particular latent print is in fact the print of a particular person. The defendants will be permitted to present their own fingerprint experts to counter the government’s fingerprint testimony, but defense experts will also be precluded from presenting “evaluation” testimony. Government counsel and defense counsel will, in closing arguments, be free to argue to the jury that, on the basis of the jury’s observation of a particular latent print and a particular rolled print, the jury may find the existence, or the non-existence, of a match between the prints.\footnote{\textit{Llera Plaza \textit{I}}, 179 F. Supp. 2d at 516. Note that while Judge Pollak was the first judge to take this compromise approach to the admissibility of latent fingerprint evidence in particular, he was following an approach taken by several other courts in the context of expert evidence on handwriting identification evidence. The first judge to take this approach was Judge Matsch, in the trial of Timothy McVeigh for the Oklahoma City bombing. No written opinion was issued, but the transcript of the discussion of this issue is available at Pre-Trial Transcr., United States v. McVeigh, 1997 WL 47724 (D. Colo. Feb. 5, 1997). Judge Nancy Gertner was the first judge to issue a written opinion in support of this approach in United States v. Hines, 55 F. Supp. 2d 62 (D. Mass. 1999). See infra notes 137-138 and accompanying text.}

There are several aspects to note regarding \textit{Llera Plaza \textit{I}}’s logic. First, the court clearly and decisively recognizes that the latent fingerprint identification evidence did not fare well under \textit{Daubert}’s strictures.\footnote{\textit{Id.}} But then, because some part of the evidence is “descriptive, not judgmental” and in partial deference to the long judicial acceptance of this form of proof, the judge determined that outright exclusion would be “unwarrantably heavy-handed.”\footnote{\textit{Id.}} His compromise was to permit the expert to show the jury the similarities in the prints at issue, to point their attention toward the data that the expert would typically use to derive his or her conclusion, but to prohibit the expert from actually providing that conclusion or opinion to the jury. The idea is that the expert would merely be showing “objective” data to the jury; it would then be up to the jury, without overt expert assistance, to decide what
meaning to give to the data, and what conclusion, if any, to reach about the likelihood that the fingerprints it had been shown derived from a common source.

While other courts have not yet taken this approach to fingerprint evidence (and of course Pollak himself reversed course in *Llera Plaza II*), a number of other judges have adopted this same structural compromise—description without ultimate conclusion—in cases involving other kinds of pattern identification evidence, notably handwriting and firearms identification.\footnote{See, e.g., United States v. Hines, 55 F. Supp. 2d 62 (D. Mass. 1999) (involving handwriting identification); United States v. Green, 405 F. Supp. 2d 104 (D. Mass. 2005) (involving firearms identification); United States v. Glynn, 578 F. Supp. 2d 567 (S.D.N.Y. 2008) (permitting a firearms identification to be deemed “more likely than not” but not permitting individualization in absolute terms). There are several other handwriting identification cases that have followed *Hines* in taking this approach. See generally Risinger, *Cases Involving*, supra note 8.}

At first glance, this compromise approach seems to be a clever and appropriate strategy.\footnote{Indeed, several academic commentators have noted it with approval. See, e.g., Michael J. Saks, *Protecting Factfinders From Being Overly Misled, While Still Admitting Weakly Supported Forensic Science into Evidence*, 43 TULSA L. REV. 609 (2007); Robert P. Mosteller, *Finding the Golden Mean with Daubert: An Elusive, Perhaps Impossible, Goal*, 52 VILL. L. REV. 723, 760-62 (2007).} Merely pointing out the similarities and differences in two visible patterns without providing any conclusion does, it seems to me, reduce or even eliminate the *Daubert* problems with the evidence. To be sure, it is not altogether clear that an expert’s testimony is even necessary to point out similarities and differences in two visual exemplars, as the jury members can look at the images for themselves and thus have access to the same visual data with or without the expert’s testimony. However, due to his or her training and experience, an expert may well be better at seeing those similarities and differences. Lay jurors may therefore be meaningfully assisted in their own observations and examinations of the visual stimuli by having the expert point out precisely what is worth looking at, how to look, and how to see minutiae, and both the similarities and differences in the visual exemplars, for themselves.

The experts, in other words, may, under this approach, provide the factfinder with a kind of educative expertise—the expert does not ask for deference to his or her authority, but rather, teaches the jury members how to see the patterns present in the fingerprint or bullet or handwriting sample for
themselves. It is as if the expert is teaching the jury how to read music, instead of playing the notes on the piano himself. The expert provides a lesson in “how to see,” and the jury then exercises its own vision and reaches its own conclusion. Framed in this way, it is fair to say that the expert assistance in pointing out similarities and any difference meets the “helpfulness” requirement for expert testimony under Daubert.

Furthermore, the lack of knowledge about error rates, the non-existence of objective standards for determining a match, and the lack of statistical models for determining the probabilities of a match do not seem nearly as problematic under Daubert, at least at first glance, when no conclusion about the meaning of the match is being provided. As Judge Gertner explained in a case involving handwriting identification,

[The expert’s] account of what is similar or not similar in the handwriting of [the defendant] and the robber can be understood and evaluated by the jury. . . . This is not rocket science, or higher math. Her conclusion of authorship, however, has a difference resonance: “Out of all of my experience, and training, I am saying that he is the one, the very author.” That leap may not at all be justified by the underlying data; and in the context of this case, is extraordinarily prejudicial.

If the expert confines herself to pointing out similarities and differences to the factfinder, while eschewing all conclusions, then the questions regarding the legitimate strength of the expert’s conclusions, or the lack of a statistical model to justify a claim about the probabilities associated with a match, or the appropriate error rates associated with the expert’s conclusions, all become moot. No match, and, indeed, no ultimate conclusion regarding identification or its absence, is being introduced into evidence.

Perhaps, then, this compromise provides an exemplary way to navigate away from the awkward spot in which the forensic sciences now find themselves? Certainly, the small handful of judges that have put forward this compromise view are to be commended, for they are wrestling valiantly with a


135 Daubert, 509 U.S. 579 at 591.

136 Hines, 55 F. Supp. 2d at 69.
set of hard questions about how to handle these powerful but inadequately tested forms of proof. It may well be that the Solomonic compromise is the best practical alternative we can come up with at present if courts deem exclusion too draconian a remedy. This compromise eliminates the excessively strong and presently unjustified claims about the strength of forensic experts’ conclusions. It represents an effort to acknowledge, on the one hand, the likely power of this evidence, while, on the other hand, not treating the evidence as if it passes Daubert with flying colors like Havard.

And yet, just like King Solomon’s proposal to divide the baby between the two women claiming to be its mother, this compromise approach is, unfortunately, fundamentally unsatisfying. Why so? Upon careful analysis, there are two significant difficulties that arise. First, it may be substantially more difficult, as a practical matter, to eliminate the evaluative aspects from testimony than one might expect. For example, Judge Gertner wrote in Hines, “The witness can be cross examined, as she was, about why this difference was not considered consequential, while this difference was, and the jury can draw their own conclusions.” But notice what the witness would be testifying to here. While she would not be testifying to an ultimate conclusion about authorship, she would be testifying as to why some differences are considered consequential and others are not, rather than simply presenting the visual content of the handwriting exemplars and pointing out similarities and/or differences. This is, already, evaluative rather than merely descriptive. She would be saying, more or less, “this kind of difference is still consistent with the two samples coming from the same person’s writing, while this other difference is not something we would expect to see in two samples of writing from the same person.” Perhaps the expert’s inferences regarding the interpretive

---

137 Id.

138 Moreover, as I pointed out some time ago, taken to the extreme, this would result in little more than a semantic difference between this “restricted” approach and permitting conclusion testimony under Daubert. If an expert can say to the jury, “Look at these seventeen meaningful similarities, and here’s why they are really meaningful” and, “Look at these three apparent differences, and here’s why I don’t think they count for anything,” the expert’s own conclusion about authorship or identification would be completely clear to the jury, even if the expert did not actually use the words “match” anywhere in his or her testimony. See generally Mnookin, Scripting Expertise, supra note 1; accord, Risinger, Cases Involving, supra note 8, at 510; D. Michael Risinger & Jeffrey Loop, Three Card Monte, Monty Hall, Modus Operandi and “Offender Profiling”, 24 CARDOZO L. REV. 193, 209-10 (2002).
meaning of this data are correct, and perhaps not. The point is
that either way, this is not merely showing the jury the
existence of similarities and differences but is, already,
evaluating their meaning, determining which differences truly
“count” in favor of a conclusion of authorship and which may
not. And thus we need again to ask a set of already familiar
questions: What data shows that the expert can do what she or
he claims to be able to do? Is there data to support the notion
that a certain kind of difference is consistent with authorship
and another kind is not? How accurate are such conclusions?
What is their error rate?

We have returned, in essence, straight back to the
Daubert problem, albeit having drilled down the analysis to a
more particularized, local, level, rather than asking validity
questions with respect to an overall conclusion about
authorship of the writing. Similarly, in the case of fingerprint
identification, consider the expert’s belief that a particular
apparent visual difference between the two prints is an artifact
rather than an actual discrepancy. If the expert is permitted to
testify that he or she believes that a particular difference is not
a true discrepancy, that too raises all the same questions about
the extent to which data supports the ability of an expert to
accurately distinguish between source discrepancies and
differences that are not interpretively meaningful, “real”
differences in the source prints versus those differences that
derive from the process of taking and making the images. At
present, adequate published data regarding the reliability and
scientific validity of these mid-level inferences is largely non-
existent.

This analysis might suggest a straightforward answer:
perhaps what is needed is simply to take a more complete end-
run around Daubert by carefully prohibiting what we might
call “intermediate” inferences about the evidence. Judges could
permit experts truly to testify only to what they see, literally
just to point the factfinder toward details on the bullet, the
print, or the handwriting exemplar that the layperson might
not have noticed without expert assistance. If experts truly
limited themselves only to description,\footnote{I am here ignoring the question of whether there is actually any such thing
as description without interpretation—my own view would be that there is not. Nonetheless, these are matters of degree, and prohibiting all inferences about the likely significance of the minutiae observed does get closer to being “just” description than does permitting the expert’s overtly evaluative statements.}

\footnote{I am here ignoring the question of whether there is actually any such thing as description without interpretation—my own view would be that there is not. Nonetheless, these are matters of degree, and prohibiting all inferences about the likely significance of the minutiae observed does get closer to being “just” description than does permitting the expert’s overtly evaluative statements.}
eliminate the lurking validity problems associated with mid-level evaluations and conclusions? Leave it entirely to the jury to decide whether particular visual differences ought or ought not to be taken as meaningful discrepancies in two fingerprints. Leave it to the jury to decide whether the quantity and types of differences in two handwriting samples suggest distinct authorship or remain consistent with one person having some degree of inevitable variation in how they form their letters.

If we truly limited the expert to “pure” description, this might indeed solve the lurking Daubert issues, though it would do so by radically curtailing the expert’s role. It should be noted that this would also be quite an unusual approach to expertise—part of what is special about experts’ roles under the Federal Rules of Evidence is precisely that experts are typically given significantly more leeway than lay witnesses to provide their opinions and conclusions. Placing forensic experts into a kind of Lockean straitjacket—permitting them only to testify to what they can empirically observe, rather than allowing them to share the inferences and judgments they make about what they see—is, in a sense, to de-authorize them as experts. Much of what a forensic expert thinks of as his or her expertise—which is precisely the experience-based ability to assess and analyze the image; to differentiate signal from noise and artifact from discrepancy; and to evaluate whether two patterns did or did not come from the same source—would no longer be permitted. This is not necessarily a bad outcome—indeed, it might well spur an increased interest among forensic scientists themselves in promoting and participating in the research that would permit the courts to grant them a more significant evaluative role. But it is important to recognize that if we take this compromise seriously, the permitted testimony must be quite limited, significantly more curtailed than it was, for example, in Hines itself.

Moreover, we should recognize that juries will likely still find such extremely curtailed evidence quite probative. Juries walk in the door with prior, culturally-based views about evidence and what kinds of evidence count, and most jurors probably arrive with a deeply-held belief in the significant power of pattern identification evidence, just like my seatmate

---

140 See generally Fed. R. Evid. 701-02.
on the airplane whose beliefs about forensic science derived mostly from her television-viewing. So even if all an expert did was point out similarities, with absolutely no testimony whatsoever about the meaning of these similarities, a jury might be quite prepared to presume that those similarities imply that the two impressions come from a common source.

This leads to the second, more serious difficulty. The still-larger problem with the Solomonic compromise is that if experts cannot provide some data-based, research-justified evidence supporting their conclusions, it is not clear that the impression evidence should be admissible at all. The problem, in a sense, migrates: it stops being a problem relating to the adequacy of the evidence supporting scientific validity, and becomes a problem of relevance and probative value.

To put the point in terms of the Federal Rules of Evidence, radically curtailing the experts’ testimony solves the Rule 702 expert evidence problem, but at the cost of creating a Rule 401/403 problem. Rule 401 requires that all evidence be relevant, and the Federal Rules define as relevant that evidence which makes “the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.”\footnote{Fed. R. Evid. 401.} Rule 403 permits the exclusion of relevant evidence when “its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury.”\footnote{Fed. R. Evid. 403.}

The problem, in a nutshell, is that if the expert cannot provide the factfinder with any admissible evidence about the meaning of the visual similarities and differences, then the factfinder has no rational basis for assessing the probative value of these observations. Certainly the factfinder can look at the visual stimuli herself, but what legitimate basis does she have for making an inference about probative value given whatever quantum of similarities she observes? The key question about the meaning of any pattern identification evidence is how much support it provides for the inference that two patterns do or do not share a common source. How often would we expect to see any given degree of similarity from two bullets that did not come from the same gun, or from two
fingerprints that were not actually impressions from the same finger?

To answer this question, juries could draw on two potential sources of knowledge: information provided by the evidence presented to them at trial via an expert, or their own experience. But in the Solomonic compromise, we are prohibiting the first kind of information, because it does not at present meet Daubert’s strictures. So juries are left entirely to their own resources and devices—only their own experience can help them assess the probative value of the patterns at issue. But a serious problem arises because—with the possible exception of handwriting identification—juries simply do not have any meaningful experience on which to draw for these conclusions. Non-experts are not in the habit of looking closely at ridge minutiae on fingerprints to develop intuitions about how much similarity might exist on the tips of two different individuals’ fingers. Ordinary people do not encounter bullet striations in their regular life. Jurors therefore have literally no personal, experience-based information that would provide any rational basis for evaluating the similarities and differences that the expert helped them to notice. While the evidence might nonetheless squeak by Rule 401, given the very low threshold for defining relevance under the Rules of Evidence, the lack of any rational basis for assessing its meaning makes the evidence both prejudicial and potentially misleading, thus rendering it excludable under Rule 403. One can think of the issue like this: because we believe forensic expert testimony often has probative value, it is therefore relevant—but because we really do not have any rational way to assess that probative value, its admissibility seriously risks being both misleading and prejudicial.

To be sure, we regularly permit jurors to engage in lay assessments of “frequencies” and to determine for themselves probative value in the trial process. If a victim who saw the perpetrator testifies that the person who assaulted her had grey hair, a beard, and a tattoo of a purple parrot on his forearm, and the defendant also has those characteristics, we

---

144 Certainly it is a reasonable inference to say that two prints or two bullets that have many similarities to one another are more likely to come from a common source than two that do not. For this reason, the evidence does, I suppose, meet Federal Rule of Evidence 401’s low standard, which deems relevant that evidence which has “any tendency” to make a matter more or less probative standard. The problem that we know it has some tendency, but cannot presently quantify how much or how little power it really has.
do not require any data to be presented to the jury on the frequency with which any or all of these characteristics actually exist within the relevant suspect population. Nor do we require any formalized model for assessing the likelihood that these traits would co-exist within a given individual within that population, or whether the different characteristics tend to be statistically independent from one another or not. We let the jury decide for itself precisely how much power and how much probative value to give to the evidence introduced that shows that the perpetrator and the defendants share certain characteristics in common. The fact that individual juror’s subjective assessments of the frequency of purple parrot tattoos might be wildly off-base does not render the evidence inadmissible.

Why, then, is the evidence of the similarities in bullet striations or ridge detail on fingerprints any different from the evidence of the purple parrot tattoo? Yes, juror assessments of the evidentiary power of the pattern similarities in a fingerprint might be substantially inaccurate, but that risk exists in the tattoo scenario as well. The difference, I would suggest, is that we believe that jurors’ ordinary lives provide them with some legitimate basis—albeit partial and imperfect—for assessing the frequency of purple parrot tattoos on people’s forearms. In ordinary life, people see each other’s forearms—and each other’s tattoos—with some regularity. Most of us probably do not go around counting how many people have tattoos, or keeping track of how many birds we have ever seen tattooed on other people’s bodies. But nonetheless, our individual paths navigating through the crowded world do give us an experiential basis for having a rough empirical sense of just how rare or common purple parrot tattoos might be. Note, in a sense, that our individual

145 Indeed, in the famous case of People v. Collins, 438 P.2d 33 (Cal. 1968), a case that came to stand for the reluctance of courts to frame questions in overtly probabilistic terms, the effort to use the “product” rule to figure out a combined probability of a variety of specific characteristics—ranging from the fact that the couple was interracial to the fact that she had blonde hair and a ponytail and he had a beard—was strongly criticized. For the classic critique of excessive confidence in probabilistic thinking in court, see generally Lawrence H. Tribe, Trial by Mathematics: Precision and Ritual in the Legal Process, 84 HARV. L. REV. 1329 (1971). However, one of the (numerous) problems in Collins was that the numbers presented to the jurors were purely speculative—they were simply the expert’s illustration of possible numbers to illustrate how the product rule worked, rather than numbers based on empirical reality. Collins, 438 P.2d 33.

146 There is a stronger argument for taking the “compromise” approach in handwriting identification cases, precisely because our ordinary lives do give us some
relationships to the evaluation of parrot tattoos are rather like the forensic expert’s relationship to the evaluation of patterns—they are experience-based, non-quantitative, and without any known error rate. The difference however, is in role: we expect juries to bring their common knowledge to their role as factfinder and to their assessment of the evidence, and we do not require this common knowledge to be validated or scientific. The same cannot be said for the expert required to meet the strictures of *Daubert*.

In addition, the strong cultural belief in forensic science evidence, based first and foremost on its roughly one hundred years of courtroom use, likely means that even without any actual experience of their own looking at comparable patterns, jurors will believe they do know the meaning of the similarities they are taught by the expert to observe. With latent fingerprint examination, the tradition of experts individualizing in court—without adequate data or research to support their claim—may well mean that when a jury is presented with latent fingerprint identification evidence without any expert conclusions, it will nonetheless believe itself quite capable of drawing a conclusion of identity. But this confidence and belief does not derive from the first-hand experience of jury members looking with care at fingerprints or bullet striations. Rather, it would stem from their prior belief that latent fingerprints can individualize, and hence if these prints appear reasonably similar, then they probably did come from the same person. They may not ever have looked at fingerprints for themselves, but from courtroom dramas on television, from criminal cases they have read about in the newspaper, from *CSI*, from their myriad cultural experiences in the world, they may well believe not only that everyone’s fingerprints are different, but also that two similar prints must necessarily have come from the same finger.\(^{147}\)

\(^{147}\) The cultural mythology of fingerprints remains deep and widespread; even my seven-year-old son told me the other day that everyone’s fingerprints are different.
Notice the concerning kind of feedback loop occurring here. The substance of the excluded evidence—conclusions about the common source of the latent print and the exemplar—is likely to be presumed by the jury precisely because that now-excluded evidence has generally been admitted in the past. If my hunch about likely jury expectations and prior understanding is correct, then the common knowledge that will be doing a good deal of interpretive work for the jury when they try to assess the fingerprint is the very evidence that was excluded. That certainly raises a significant Rule 403 problem.

Therefore, unless the expert can provide the jury with some valuable, data-driven basis for interpreting the meaning of the similarities and differences that are presented, the pattern identification evidence presented under the Solomonic compromise still ought to raise significant admissibility concerns. This is not because the experts’ descriptions are insufficiently reliable under Rule 702, but because of the lack of any basis—apart from the jury’s likely prior belief in the very conclusions that have been excluded—for meaningfully evaluating the probative value of the evidence.

The Solomonic compromise is thus far more problematic than it initially appears. However, I recognize that it might nonetheless be a reasonable second-best solution given the present impasse, and I therefore do believe it has a legitimate place in the judicial “toolkit” for assessing pattern identification evidence. Moreover, there might be ways, at trial, to dislodge, at least partially, whatever prior beliefs the jurors had about the basis for individualization or the meaning of a “match.” If the Solomonic compromise were coupled with effective evidence to demonstrate that there is not a statistical basis for reaching a conclusion about whether the two patterns come from a common source, nor any validated metrics for evaluating how much similarity is needed to warrant such a conclusion, the jury might call into question its prior assumption that similarity necessarily meant identity.

There is little doubt that evidence of similarity of bullet striation patterns or handwriting similarities or friction-ridge patterns on fingerprints often does have probative value. The question is how much. The hard question facing judges is what to do given that the answer to that “how much” question is, “we really don’t quite know.” It seems clear that on the basis of current empirical knowledge, judges should absolutely not permit conclusions of individualization to be made by experts,
and softening an individualization conclusion by framing it as “opinion” rather than “fact” does not enhance either its validity or its research basis. Nor should either judges or experts kid themselves that adding in some fudge words like “to a reasonable certainty” changes the analysis at all.\(^{144}\)

The harder question is whether the fact that we don’t know precisely how much probative value to assign to evidence of any given quantum of similarity—and that a jury assessing the meaning of these similarities likely has no first-hand life experience to help assess probative value—ought to lead to the evidence’s exclusion. As a matter of logic, without some meaningful basis for assessing the probative value, the evidence becomes literally uninterpretable—and this should rationally argue in favor of exclusion. But given that experts’ experience and our collective cultural experience over the last century with fingerprint identification evidence supports the inference that the probative value of this evidence is likely to be quite substantial, the Solomonic compromise, notwithstanding its awkwardness, might be as good an option as any for the moment. The Solomonic compromise, is in a sense, an approach based on a bet that the cultural belief in forensic science will, when more research has been conducted, largely prove to have been warranted. The more confident a judge is that future research will be likely to validate the claims fingerprint examiners have been making in court for the last 100 years, the less troubled she might be by letting those conclusions in through the back door, via jurors’ prior beliefs about forensic science and its credibility. Put like that, the Solomonic compromise becomes a more nuanced version of judicial ACE-V—all things considered, it’s valid enough to permit, so long as we force it into an uncomfortable straitjacket, a straitjacket that cannot be expected to stay on within the jury room.

Thus, even though it does have an appropriate role in the judicial toolkit, the Solomonic compromise is, at best, an awkward doctrinal solution. If only we had the data to support them, it would be far better to permit expert inferences about the meaning of similarities and conclusions about probative value. It is also worth noting that outright exclusion would put significantly more pressure on the forensic science community to cooperate with and to lobby for the research that would

---

\(^{144}\) Taylor, 663 F. Supp. 2d at 1179.
bring them back into the evidentiary fold. Nonetheless, I can certainly understand the attraction of the Solomonic compromise as a second-best solution for probative but underresearched evidence that has a long history of legal acceptance.

IV. IN DEFENSE OF EXCLUSION (FOR NOW)

The other viable alternative is, obviously, outright exclusion, at least in some circumstances. Given judges’ treatment of forensic science evidence so far, arguing for exclusion of this evidence can feel, from a practical standpoint, more than a little quixotic.

And yet, the truth of the matter is that at present, pattern identification evidence does not have the empirical data to back up the claims made in court. Moreover, just as with DNA evidence—which after an initial honeymoon period, was excluded by a number of jurisdictions for a short period of time because of concerns about the subjectivity of standards for determining a match; insufficient research into the underlying questions of population genetics; and general technical sloppiness—exclusion would be a great motivator for pursuing the research necessary to justify admissibility.

Moreover, I want to suggest that the use of the exclusion option—which I do think should have a central place in the judicial toolkit, given the present lack of an adequate research basis supporting validity—could also be quite short-term in most pattern identification arenas. For as I will suggest, what ought to be the minimum necessary information to establish adequate validity under Daubert is simply not that onerous. Good proficiency tests, which show the extent to which examiners make errors in a variety of different levels of difficulty, should suffice to support a finding of adequate

149 I do not explicitly address the question of when courts should exclude, and when the Solomonic approach is warranted. Partly that is because I think, in many instances, either approach could be a legitimate exercise of judicial discretion given the current state of our knowledge. Moreover, the choice ought to be informed by the particulars—were I a judge ruling on admissibility, I would be far more likely to exclude a single AFIS-generated match than evidence that linked multiple, high-quality prints to several different fingers of the same individual. Even without formal metrics for sufficiency, common sense—and Kumho Tire—tell us that all identification tasks are not created equal, and the judicial response can, and sometimes should, vary as a result.

validity, presuming that the error rates discovered through this testing process are tolerably low, and the match between what was tested and the “task at hand” in the particular case is sufficiently close.

More generally, as I have also argued elsewhere, I want to suggest that under *Daubert* it is the ‘testing’ criteria that should matter most.\(^{151}\) I want to make a distinction between explanation or description, on the one hand, and testing on the other. Explanation of the methods and descriptions of the processes used by an expert should not be permitted to substitute for adequate testing of validity. It is this mistake, I believe, that has often plagued the courts when evaluating forensic science. Judges hear about the ACE-V process and they listen to examiners describing their approach, and judges are persuaded, it seems, that this methodology therefore works.\(^{152}\) They, like the factfinders, see the similarities magnified and put up on a giant chart, and they “see” the method in action, and believe its power. Even putting aside that ACE-V is not, in fact, much of a specified methodology, the more important point is that the courts should care less about the details of the method at issue or its seeming plausibility, and more about what evidence there is to support the conclusion that the methods actually work.

In other words, judges have been lulled by plausible descriptions and seemingly persuasive explanations of forensic science techniques into dismissing the importance of the nearly complete lack of empirical support for the experts’ claims. All of the opinions discussed above as examples of legal “ACE-V” have this quality. These judges recognized the lack of testing, but found enough within the description of the method that seemed credible that they decided that the technique passed muster.\(^{153}\) This, in my view, is a mistake.

To be sure, the pattern identification techniques of forensic science do have a certain intuitive plausibility to them, and their early acceptance was linked, in no small part, to this cultural plausibility.\(^{154}\) We all have some experience identifying handwriting for example: no doubt most of us believe that we


\(^{152}\) See supra Part III.B.

\(^{153}\) See generally supra notes 105-113 and accompanying text.

\(^{154}\) See generally Mnookin, *Fingerprint Evidence*, supra note 1; Mnookin, *Scripting Expertise*, supra note 1.
could likely distinguish our mother’s handwriting from that of our closest friend. We can stare down at our fingerprints as well, and can even see for ourselves how the ridges and whorls on our fingers vary from those of others. Moreover, these patterns have a particular quality—they are semi-legible in that they can be seen and their differences can be noted by all of us, not just by experts, while at the same time, they still do require significant expert analysis and interpretation.

But the inherent plausibility and the semi-legible quality of these materials combine to make it particularly easy for judges to be seduced by description and explanation into failing to ask what they ought to be asking both as a matter of doctrine and as a matter of logic. Their focus should be on the degree of empirical support for the actual, specific claims being made. They should be asking precisely what evidence supports this particular evidentiary claim. With latent fingerprint evidence, for example, the most central question ought to be: how accurate are examiners when matching latent prints to a particular source; latent prints which are often partial, frequently smudged, and perhaps even distorted? For any of the forensic sciences, what judges ought to ask under Daubert is precisely this: what empirical support shows that the expert can actually do what she claims to do? What data, what testing, would be necessary to justify the claims being made in the expert’s testimony?

An important corollary of this focus on testing is that it ought to be permissible under Daubert for the technique or method to be a kind of “black box.” A “black box” is a technique or method that we do not necessarily understand, but that we can nonetheless test to see what it does and how it works. If there is sufficient testing to show us that it works, I do not believe that the proffering party should be required under Daubert to show how it works. My argument is that it is far less important to pry open this black box than it is to ask whether the technique has been tested under conditions similar to those at issue in court. Peering inside the black box to see how it works is less critical for an assessment of validity than assessing whether input/output testing shows that it works. Indeed, peering inside, on its own, ought to be deemed neither necessary, nor sufficient under Daubert.
To make the point more concretely, when it comes to fingerprint evidence, for example, what we really need in order to justify admissibility is, at a minimum, some very good proficiency tests to show us what experts can do and to gain information about how often they make mistakes. These proficiency tests ought to be appropriately difficult and should mirror the range of difficulty found in actual casework. In addition, and critically, they should include some of what fingerprint experts would call “tough idents”—prints that are particularly difficult to identify. These proficiency tests should ideally be part of the normal stream of casework, so the examiner doesn’t know she is being tested and therefore possibly perform her analysis with a greater degree of care than usual.

Recall that one of the research lacunae with respect to fingerprint evidence is that we do not yet have an operational statistical model of fingerprints, a model that could provide us with empirically grounded information about the likelihood that two prints selected at random would both have a specified set of minutiae. Recall also that ACE-V is extremely vague, and does not come close to providing a fully developed and adequately articulated method with detailed specifications. Recall, in addition, that interpretation of fingerprints is subjective, without shared norms or rules about what is required. My argument is that these gaps in our knowledge base, though unfortunate, are not fatal. More precisely, I want to suggest that neither a detailed specification of method, nor statistical validation of frequencies ought to be seen as necessary criteria for using fingerprints as evidence in court under Daubert’s strictures.

In other words, it ought to be acceptable for latent print examiners themselves to operate as a kind of black box—for the

---

note 1, at 40-43. To be sure, I am largely skating around hard questions about whether testing can ever be wholly “black box,” or what some degree of internal understanding of the method is necessary in order to design good black box tests, and if so, how much. Still, the key point is to urge courts to focus not on asking experts to explain what they do, but rather to show that what they do works.

157 There is not presently any validated metric of difficulty for fingerprint comparisons. Nonetheless, examiners do have at least informal understandings of what makes particular comparisons more or less difficult.

158 There is presently a “black box” study of more or less this sort underway, being conducted by Austin Hickland through the FBI. It sounds, from the descriptions, like an extremely promising project. However, at this point neither the research protocol nor the results have been made public.

159 Haber & Haber, supra note 1.
examiners themselves to be the instrument, the technique. If we know through proficiency tests that they get the right answers a very high proportion of the time over a range of circumstances that mimic what they encounter in actual cases, their testimony ought to be admissible under Daubert. Even if we do not fully know how their method operates—indeed, even if they do not fully understand it themselves—this ought not to prevent us from making evidentiary use of their conclusions if we have sufficient information about how accurate they are and what circumstances and conditions seem to increase the risk of error.

Let me make one thing clear. I am certainly not opposed to research and inquiry that peers inside the black box to learn more about how the methodology works and aims to improve it. There are some impressive efforts underway to model fingerprint evidence statistically, and this is extremely important research that I fully support.\textsuperscript{160} Of course, fingerprint experts should continue to hone their methods, continue to work for better understanding of how they could improve their processes, and if researchers can help practicing forensic scientists to develop validated standards for interpretation to increase objectivity, that will be all to the good. All of these efforts at better understanding, formalizing methods, and improving practices—forms of opening up the black box and peering inside—ought to be welcomed, celebrated, and encouraged, and perhaps most importantly, funded.

They just ought not to be necessary requirements for admissibility in court. Nor, standing alone, should they be sufficient, unless they are also accompanied by evidence that these methods, in actual practice rather than in theory, truly work. To make the argument by analogy: Which would make a driver more comfortable—knowing that there was substantial theoretical knowledge suggesting that the brakes on her vehicle ought to work, because of a great deal of scientific study of the theoretical mechanisms underlying the brakes—or a substantial quantity of actual testing of identical vehicles showing that the brakes have worked consistently in conditions that mirror her situation? While both forms of knowledge are valuable, I would posit that the actual, on-the-ground testing, quite appropriately gives us a good deal more comfort than the theoretical knowledge standing alone.

\textsuperscript{160} Neumann et al, supra note 31.
I am arguing therefore, that the courts’ central focus, especially for a kind of evidence that does not presently have a formalized method, should be on testing. Courts should also insist upon a close relationship between the testing that has been done and the claims that an expert makes. This is what Daubert calls the question of “fit” and it is also part of what Kumho Tire emphasizes as the need to focus on “the task at hand.”161 Strong claims should require strong tests to back them up. For example, fingerprint examiners claim to be able to individualize—to match a print to one unique source, to one finger on one person out of everyone on earth who has ever lived or who will ever live.162 There is, as I have already stated, not yet sufficient validation to support that claim. Nor do I think that proficiency tests of the sort that I am describing and suggesting as a prerequisite to admissibility would be sufficient to support this claimed ability to individualize. Even with excellent proficiency tests, experts would therefore need to modulate and moderate the strength of their conclusions to some degree. But excellent performance on difficult proficiency tests might be sufficient to support a conclusion of a softer sort, a conclusion such as, “Based on my knowledge, testing and experience, I would not expect to see this degree of similarity between two prints unless they came from a common source.” This is still, in a way, fudging—for without a working statistical model, we cannot have a robust and quantified sense of the likelihood two prints with a given degree of similarity might have come from different sources. But if we knew that examiners in general—and that examiner in particular—had succeeded in proficiency tests that required her to make difficult identifications and equally difficult exclusions, we could legitimately believe that her knowledge and experience did adequately support her opinion that they came from a common source.

A signal advantage of focusing on testing rather than explanation is that it has the virtue of being manageable, from a practical point of view. There are simply no insurmountable obstacles to developing appropriately difficult proficiency tests. The roadblocks to doing it thus far have been cultural and institutional, not scientific. By contrast, developing a valid statistical model of fingerprint evidence is a daunting task.

161 Daubert, 509 U.S. 579 at 591.
162 See supra note 35 and accompanying text.
Even with significant research, it likely is years away, perhaps even decades. By contrast, proficiency test development could be done quite rapidly. And I have little doubt that if courts began to exclude fingerprint evidence for the lack of such tests, the requisite testing mechanisms would be developed in extremely short order.

Implicit in my argument here for a focus on testing is a quasi-“best evidence” approach to the evaluation of expert testimony. I do not believe that Daubert should be understood as requiring a fixed and unchanging amount of evidence in any particular area. Rather, the question ought to be whether the proffered expert evidence is as reliable as it can reasonably be, considering the context and circumstances. Validity under Daubert should not be understood as an on/off switch, or as an all-or-nothing proposition, in which items of evidence are inherently reliable or unreliable. The question for the court is whether they are reliable enough—and this depends both on what inferences are sought to be drawn from them, and partly on whether the evidence offered was as reliable as possible under the circumstances. To put it differently, to pass muster under Daubert, the judge must have some legitimate justification for believing that the evidence is sufficiently reliable that a jury should hear it. And the evidence of validity should be as strong as it reasonably can be, given the circumstances. It is, therefore, partly because proficiency tests are genuinely ‘do-able’ that courts ought to require them. Concomitantly, courts should hesitate before finding that studies or research that are beyond current scientific capacity are nonetheless required under Daubert.

In other words, if an expert—or an entire field—has done as much as can reasonably be done to establish validity, and our still-imperfect information suggests significant probative value, excluding the evidence because our validation knowledge is incomplete is not likely to be justice-enhancing—especially if our knowledge includes reasonable estimates of error rate, so that the factfinder can, at least in theory, adjust its assessment of probative value accordingly. But by the same token, if an expert—or a field—fails to undertake those tasks which could reasonably be done to establish validity, and

---

instead simply tells the court, “Trust me!” the court should balk at the lack of data. Under this approach, evidentiary assessment will necessarily be dynamic. A reasonable amount of research at Time A may not continue to be adequate at Time B, as time passes, and the potential for developing relevant data on the salient questions increases.

It is worth noting that what drives me to a ‘better evidence’ principle is very much the same kinds of concerns that have motivated Margaret Berger in her important work on causation in tort law. Professor Berger has been acutely concerned about the difficulties plaintiffs have faced in establishing causation in toxic torts cases.164 Throughout her career, she has been sensitive to the significant difficulties that scientific uncertainty poses for our legal system. In the torts context, she has, therefore, searched for ways to protect plaintiffs’ interests while simultaneously respecting the need for high-quality scientific information within our system of adjudication. Professor Berger has offered some insightful and creative solutions to the tension between the need for good scientific information, and the need to recognize that Daubert (at least as interpreted by courts in the pharmaceutical torts context) often expects too much. She has therefore searched for principled ways around general causation, sometimes through the possibility of expanding other kinds of tort claims—like the right to informed choice in making a decision about medication, or the duty of a company to keep itself informed of risks through the pursuit of reasonable research.165

By contrast, in the forensic science challenges, courts have interpreted Daubert as to expect rather too little, instead of too much. But the underlying concerns—how do we deal with uncertain knowledge, and how can we generate the right incentives within the legal system both to do justice and to produce better information—are strikingly similar. A ‘better evidence’ principle may be our best bet—certainly in the forensic science context, but perhaps more generally across the board as well, though its potential application in other expert contexts goes entirely beyond the scope of this article.

Neither a focus on testing nor a more general ‘best evidence’ approach to expert evidence is a panacea. There will still be difficult questions for courts, especially regarding just how much testing is required to justify admissibility. I do not pretend that answering those questions will be easy or straightforward. But they are, I think the right questions for courts to ask. They are no doubt challenging questions for many judges, most of whom typically lack much background in science, and who perhaps went to law school in part precisely to stay far away from such technical matters. Judges typically therefore will lack epistemic competence—the ability to evaluate the knowledge being offered the way an insider to the field would evaluate it.  

Nonetheless, I believe that by focusing on testing, courts are less likely to be misled, less likely to be lulled by compelling-sounding but as-of-yet unproven explanations, and less likely to gloss over remarkable gaps in what is known. Testing is a narrower, more tightly bounded inquiry. It explicitly directs judges to the question of what data is available to support the expert’s conclusions, and to focus on the specific fit between the data and the claim. Might judges misunderstand a study, fail to notice methodological flaws, or misconstrue what inferences can legitimately be drawn from a given research result? Absolutely. But at least these errors of interpretation would be the result of a focus on the data itself—which suggests, first and foremost, that there actually is some data on the relevant questions. Given our present state of affairs, this in itself would be a major improvement.

In addition, recall that I said at the outset of this article that we needed a two pronged solution to our problems of forensic science—both stricter scrutiny by the courts and greater federal support for research and regulatory oversight. Part of why we absolutely need the two-pronged solution is as a check on these concerns about the epistemic competence of the courts. If there is a research establishment funding work on forensic science—including the funding needed both to develop proficiency tests, and to evaluate them—it is highly likely that academic researchers will increasingly be attracted to the study of forensic science. Research follows money like bees

---

follow pollen. Fund it and they will come. And the structures of academic science and social science—with peer review, academic evaluation, as well as the review and evaluation that could be offered by a new federal agency if one ever were to be developed—could all offer a useful check on the quality of the proficiency tests that need to be developed.

There is one final advantage that would result from judges taking *Daubert* seriously in the context of pattern identification, and, therefore, either dramatically restricting or excluding evidence. At present, virtually no one believes that a new federal agency—NIFS, as suggested in the NAS report is likely. Whether the federal government will even create a new institutional space somewhere within existing federal agencies, dedicated to important issues of forensic science, is far from certain. There is a serious risk that at the agency level, nothing of genuine or transformative import will result from the NAS report.

However, if judges were to begin to take seriously the implicit lessons of the report, the odds would change dramatically. Why so? Because it would mean that a number of important forensic sciences would be likely to find themselves—only temporarily, I would expect—excluded from the courtroom due to a lack of adequate testing of the validity of their results and methods.

And if that were to begin to happen, what would ensue? My strong instinct is that the degree of interest in NIFS or at least a more modest surrogate for NIFS within an existing agency, would skyrocket. Obviously, the reason for judges to apply *Daubert* meaningfully and therefore likely exclude at least some forensic science evidence cannot primarily be to incentivize the creation of a federal agency. Judges’ appropriate focus is on the particular cases before them, and explicit reference to the external incentive effects their decisions might generate is not a first priority; even thinking in such terms may make some judges uneasy. But in this instance, there is a happy convergence: to wit, if the courts begin to do their job well, it is likely to help bring about precisely the broader public policy initiatives that are also necessary in this area.

CONCLUSION

What will the future hold for forensic science? If we look back on these methods and techniques twenty years from now, what will we see? It is perhaps entirely safe to predict that no
matter what happens, the practices of forensic science still won’t look like the television version, so my airplane seatmate is destined for ongoing disappointment with the real world. But I do not think it is utterly unrealistic to hope that the pattern identification sciences will be on a much more substantial and solid empirical footing than they are today. Ironically, they may well look somewhat less strong than they do at present. Testifying experts will certainly no longer be able to espouse an error rate of zero, and they will likely need to give up the claim that they are able to individualize. But by acknowledging their weaknesses, and honestly assessing their capacities and limitations, they will truly be far stronger than they are at present, and far more worthy of credence and respect. Will forensic science transform itself as it should? That will depend, I believe, in significant part on judges, and whether they are prepared, at long last, to evaluate pattern identification evidence with their eyes wide open and their heads out of the sand.

For all of our sake, I hope they are up to the challenge.
Revealing and Thereby Tempering the Abuses of Government-Created Evidence in Criminal Trials

Robert P. Mosteller

INTRODUCTION

I am delighted to add my contribution to this symposium in honor of the academic achievements of Professor Margaret Berger. I do this through three points. The first is Professor Berger’s commitment to providing the jury with information to do its job more effectively and her faith in that institution when properly armed with adequate information. The second is her particular remedy of requiring the recordation of evidence created by the government, and revealing that record to the jury as a way of tempering the corrupting influence of the government’s hand in the evidence development process. I heartily endorse these positions and find them to have widespread and enduring applicability. Finally, I comment on an admirable characteristic that I have found constant throughout Professor Berger’s work that adds to its brilliance: her reasoned judgment.

I anchor my comments in arguments Professor Berger made regarding the Confrontation Clause to the Sixth Amendment of the United States Constitution at a time when Ohio v. Roberts provided the controlling paradigm, before it was replaced by Crawford v. Washington. I do not wish to suggest that Professor Berger was a defender of the trustworthiness/reliability system of Roberts, for she was not.3

---

1 J. Dickson Phillips Distinguished Professor of Law, University of North Carolina School of Law. I want to thank Professors Ed Cheng and Jeff Powell for their comments on an earlier draft of this essay.
2 448 U.S. 56 (1980).
3 See Margaret A. Berger, The Deconstitutionalization of the Confrontation Clause: A Proposal for a Prosecutorial Restraint Model, 76 MINN. L. REV. 557, 605-13 (1992) (recognizing the “illusory protection afforded to a defendant by the evidentiary version of confrontation” and arguing to give greater protection against admission of
She did, however, attempt to make that system more protective of the central principles she believed the Sixth Amendment embodied.

Some of her ideas may fit within the Crawford framework as its detail is fleshed out in the future by the Supreme Court and lower courts. Some of these ideas may not, but these sound arguments are still worth noting since they may be embodied in legislation or in a new generation of procedural protections.

I. AN INSIGHTFUL AND INFLUENTIAL BRIEF IN IDAHO V. WRIGHT: PROTECTING THE DEFENDANT’S CONFRONTATION RIGHT BY PROVIDING THE JURY WITH THE BASIS TO ASSESS THE DECLARANT’S STATEMENT

My first specific example of a contribution by Professor Berger is the amicus brief she authored for the American Civil Liberties Union in Idaho v. Wright. In accord with her basic position, the Supreme Court concluded that the state had failed to demonstrate the requisite showing of reliability for a hearsay statement admitted under the catchall exception and reversed the conviction. As I stated in an earlier article, “Precisely why the Court decided to find the hearsay in Wright inadmissible because not supported by particularized indicia of reliability cannot be clearly established. However, the decision may have flowed from the arguments made by Professor Margaret Berger in an amicus brief.”

Indeed, one finds many echoes of Professor Berger’s arguments in the facts described in Justice O’Connor’s opinion in Wright that apparently led to the outcome in the case. Unfortunately as discussed in later parts of this essay, one does not find in the Court’s opinion all of her proposed solutions.

---

5 Id. at 827.
7 Professor Berger noted that “[i]ndicia of reliability are ‘easy to come by,’ [and that] . . . one part of the majority opinion in Idaho v. Wright reads like a handbook instructing prosecutors how to offer a child’s hearsay statement with requisite ‘particularized guarantees of trustworthiness.’” See Berger, supra note 3, at 606.
Wright involved testimony of Dr. John Jambura that contained the “statements” of a two-and-one-half-year-old child, Kathy Wright. These statements led to the conviction of Kathy’s mother for “lewd conduct with a minor” for allegedly assisting a male companion in raping both Kathy and her five and one-half year old sister. Kathy’s statements were admitted through Dr. Jambura, and she did not testify because the trial court found that she was “not capable of communicating to the jury.”

Professor Berger emphasized the importance of the confrontation right to the jury being able to perform its role: “The statements made by Kathy to Dr. Jambura lie at the heart of this case. Yet, for a number of reasons, the jury could not assess their reliability with any degree of confidence.” Chief among these reasons was that “it is not even clear from the doctor’s testimony precisely what words Kathy used.” The statement “is not being reported in its entirety, contains too few details to confirm its consistency with the supposed event, and was elicited in response to leading questions designed to confirm the questioner’s hypothesis.” She noted the lack of “any verbatim record of the interview,” which she would argue in a later article should be turned into a potential requirement and remedy.

Professor Berger argued that the key role of the Confrontation Clause is to enable the jury to do its job of deciding guilt and innocence in the difficult cases where the jurors’ albeit imperfect human instincts and judgments are all that stand between a just and an unjust verdict. With a child as young as Kathy, she emphasized that the jury needed to see the child testify (or at least have her exact words) so it could assess whether Kathy had reached the developmental stage where she “understood[ed] the need to tell the truth, or could distinguish fact from fantasy.” Additionally, the jury could not evaluate Kathy’s capacity for communicative speech.

---

8 See Wright, 497 U.S. at 808-12.
9 Id. at 809.
10 Brief for American Civil Liberties Union as Amicus Curiae Supporting Respondent at 9, Idaho v. Wright, 497 U.S. 805 (1990) (No. 89-260) [hereinafter Amicus Brief in Wright].
11 Id. at 10.
12 Id. at 16-17.
13 Id. at 4.
14 Id. at 9.
15 Id. at 9-10.
Berger recognized that “[f]actors such as the speed and flow of a witness’ speech, as well as articulation, intonation, mannerisms of speech and use of nonverbal modes of communication, on direct and on cross examination, enter into a jury’s assessment.” As she acknowledged, cross-examination of young children is a difficult enterprise, potentially made even more difficult by the psychological process of confabulation whereby details from imagination and earlier responses merge with the actual memory of the event to make the child erroneously believe a flawed version of the events. But it should still be provided. She summarized a number of concerns as follows:

The jury cannot evaluate accurately whether [Kathy’s] statement recounts a past event, or is the consequence of suggestive questioning in alien surroundings, in the presence of strangers, after undergoing what must have been an extremely unpleasant physical examination. In the absence of Kathy, the jury did not have the information needed to assess the appropriate weight to be given Kathy’s statement.

Professor Berger noted the special difficulty posed when the alleged statement of the child is presented through an expert, which gives her purported testimony through an impressive medium but without adequate testing. “Interposing the expert between the declarant and the jury deprives the jury of its right to make determinations of credibility.” Having Dr. Jambura on the stand was not, she argued, an adequate substitute. “Cross-examining the expert is not the equivalent of cross-examining the declarant upon whose statements the expert is relying in expressing his opinion. To the contrary, the defendant may be deprived of his rights to confrontation if he has no access to the declarant upon whom the expert is relying.

The difficulty she notes here for statements introduced through experts reverberates into the new system created by Crawford. A clear challenge yet to be addressed by the Supreme Court under the “testimonial statement” approach is whether its determination that statements not “offered for the

---

16 Id. at 10.
17 Id. at 11.
18 Id. at 17.
19 Id. at 18.
20 Id.
truth” are outside Confrontation Clause protection\textsuperscript{21} applies to statements admitted for the limited purpose of supporting the expert’s opinion. Although the answer under the hearsay rules is that such statements are not offered for the truth, the answer should be different under Confrontation Clause analysis for the straightforward reason that Professor Berger offers. Unfortunately, that difference appears to be eluding many lower courts presently as they use the wooden analysis of following the hearsay definition\textsuperscript{22} and not the importance of confrontation to enable the jury to evaluate the accuracy of the second-hand account.

Professor Berger’s arguments are simple, powerful, and correct. As part of an overall pattern of rights within the Sixth Amendment, a major role of the Confrontation Clause is to empower jurors to do their task properly and accurately. The goal is the testing of the witness’ version of events in front of the lay factfinders, part of our inherited overall system of live witnesses presenting their evidence at a public trial. Her specific vision, which I believe is sound, is not of a Confrontation Clause that provides a “get out of jail free” right but rather a guarantee that seeks to maximize the actual “confrontation” of the jury with the grist from which its members can reach their own judgments. That involves both the declarant’s direct testimony and the testing through cross-examination by counsel for the accused in front of those jurors.\textsuperscript{23}

In support of the Idaho Supreme Court,\textsuperscript{24} Professor Berger argued that the Court should require the recording of children’s statements when elicited by prosecutorial authorities in adversarial situations.\textsuperscript{25} Although her proposal was out of sync with both the Wright Court and today’s Court, the suggestion of using “prophylactic rules as the instrumental

\textsuperscript{21} See Crawford v. Washington, 541 U.S. 36, 59 n.9 (2004) (citing Tennessee v. Street, 471 U.S. 409, 414 (1985) in which a clearly testimonial statement and confession by a co-defendant that differed substantially from the defendant’s version of the events was used to refute the defendant’s claim that he was coerced into a confession that tracked the co-defendant’s statements).

\textsuperscript{22} See Julie A. Seaman, Triangulating Testimonial Hearsay: The Constitutional Boundaries of Expert Opinion Testimony, 96 GEO. L.J. 827, 846-48, 854-57 (2008) (noting lower court rulings that such statements are exempted from confrontation scrutiny and the fallacy of the reasoning).

\textsuperscript{23} See Amicus Brief in Wright, supra note 10, at 7.

\textsuperscript{24} See State v. Wright, 775 P.2d 1224, 1230 (1989) (concluding that because there was no audio or videotape of the interview, the dangers of unreliability could “never be fully assessed”).

\textsuperscript{25} See Amicus Brief in Wright, supra note 10, at 20-27.
means to further constitutional objectives" stands in excellent company with Sixth Amendment precedent and with good judgment that is being utilized today in many sectors as we recognize the need to protect the innocent.

Finally, she argued that corroboration of the statement’s truth by external evidence could not be a basis for declaring the statement trustworthy and in turn could not justify its receipt under the Roberts’ trustworthiness/reliability system. The Court reached this same conclusion largely applying hearsay theory. Professor Berger argued for it on a different and more enduring basis. It was part of her broader view that the purpose of the Confrontation Clause is to enable the jury to assess the declarant’s statement, which was part of a theoretical vision that became clear in the article to which I turn next. Her vision is that the Confrontation Clause has a particular role in protecting the accused from government developed evidence. In her brief, she articulated one ramification of that view:

If confrontation is to be excused when corroborating evidence exists, the constitutional right . . . will become meaningless. Prosecutors would be encouraged to rely on weak witnesses whom they would be able to bolster by hearsay evidence that would not violate the Confrontation Clause because it was corroborated. Such bootstrapping would spell an end to the constitutional right embodied in the clause . . . .

II. FORMULATING THE FUNDAMENTAL ARGUMENT TO CONTROL MUCH GOVERNMENT-CREATED HEARSAY BY REQUIRING THE RECORDING OF THE PROCESS OF CREATION

My second specific point of reference is Professor Berger’s article, The Deconstitutionalization of the Confrontation Clause: A Proposal for a Prosecutorial Restraint

---

26 Id. at 20.
27 See Matthew D. Thurlow, Lights, Camera, Action: Video Cameras as Tools of Justice, 23 J. MARSHALL J. COMPUTER & INFO. L. 771, 772-73 (2005) (recognizing a broad public campaign to record as a means of protecting the innocent and noting that three states through judicial decision and three others through legislative action have mandated recording interrogations).
28 See Amicus Brief in Wright, supra note 10, at 19-20.
29 See Idaho v. Wright, 497 U.S. 805, 819-21 (1990). The Court did recognize the danger of “bootstrapping” weak evidence into the case, see id. at 823, but it did not give the prominence to this danger of prosecution created evidence that Professor Berger’s prosecutorial restraint model would justify.
30 Amicus Brief in Wright, supra note 10, at 20.
In this article, she argues for special scrutiny for hearsay statements made by declarants to government agents. Although the Court did not cite her article as influencing its determination to adopt a new paradigm in *Crawford*, Justice Breyer did cite it as one of three academic articles suggesting a new approach in his statement of personal dissatisfaction with the Roberts approach prior to the *Crawford* decision.\footnote{32}

Professor Berger’s position that special attention should be paid to statements elicited by government agents is largely consistent with the Court’s new testimonial statement approach. In *Crawford*, the Court stated that “[i]nvolvelement of government officers in the production of testimony with an eye toward trial presents unique potential for prosecutorial abuse—a fact borne out time and again throughout a history with which the Framers were keenly familiar.”\footnote{33} While her remedies do not fully appear to fit the new paradigm, my point is not directed towards the mismatch. It is instead about the correctness of the central insight and the importance of her basic approach and remedy.

That approach is to give special scrutiny to the situations where government agents and the prosecution are involved during the formation of evidence and have the ability to affect its development and content. In those situations, her remedy is to require additional safeguards that would enable the jury to assess the reliability of the process and the impact of the government’s role in it.\footnote{34}

If the declarant is produced by the prosecution as a witness at trial, ordinarily no special protection is required even though the government had a role in securing the statement. However, additional protections are needed if the declarant is particularly vulnerable. For vulnerable witnesses, such as mentally unstable witnesses, children, or someone like

\footnotesize{\textsuperscript{31} Berger, supra note 3, at 605-13.}  
\footnotesize{\textsuperscript{32} See Lilly v. Virginia, 527 U.S. 116, 140 (1999) (Breyer, J., concurring).}  
\footnotesize{\textsuperscript{33} Crawford v. Washington, 541 U.S. 36, 56 n.7 (2004).}  
\footnotesize{\textsuperscript{34} Berger, supra note 3 at 561-62.}
the victim in *United States v. Owens*\(^{35}\) who was suffering from a head injury, the prosecution would be required to produce either a tape of the interview or transcript of the hearing where the statement was produced.\(^{36}\) For coconspirator statements, she draws a distinction between those made to true conspirators who are private citizens involved in crime and acting independent of the government, and those made to undercover agents and recognized informants. Because of the paramount concern of government-shaped statements, she would require exclusion of statements made to government agents and informants who, at the time of the statements, were already cooperating with the government, unless either the declarant was produced as a witness at trial or the conversation with the government agent or informant was recorded.\(^{37}\) She argues the same approach—the required recording of statements—should be used with children when the person conducting the interview is doing so at the behest of the police.\(^{38}\)

Under the testimonial statement approach, significant protection will be provided in a number of the situations

\(^{35}\) 484 U.S. 55 (1988) (involving brain damaged victim of an assault who identified the defendant as the perpetrator during a pre-trial identification procedure).

\(^{36}\) See Berger, supra note 3, at 607-08 & n.207.

\(^{37}\) See id. at 608-09.

\(^{38}\) See id. at 611-12.

I have been of two minds regarding Professor Berger's approach to videotaping children's statements. On the one hand, I took issue with what I understood to be her approach of permitting the statements to be introduced without confrontation, even if accusatory, as long as the interview was videotaped. See Robert P. Mosteller, *Remaking Confrontation Clause and Hearsay Doctrine Under the Challenge of Child Sexual Abuse Prosecutions*, 1993 U. ILL. L. REV. 691, 752 n.293 (1993) (noting substantial agreement with Professor Berger's approach but disagreeing that documenting the conversation would provide an adequate alternative). On the other, I recognized that there are very good reasons for videotaping early statements by children and that it would be unfortunate if *Crawford* caused the practice to be discontinued. See Robert P. Mosteller, *Crawford v. Washington: Encouraging and Ensuring the Confrontation of Witnesses*, 39 U. RICH. L. REV. 511, 539-40 n.144 (2005) (noting specifically the reduction of trauma from multiple interviews).

In the end, I am a firm believer in realistic solutions, and the documentation of the interview through mechanical recording is a very sound second-best solution that would provide the jury with a far better look into the circumstances and accuracy of the incriminating statements than does recital of its contents from the perspective of the government agent who obtained the statement. My preferred solution is not exclusion of the statement but both the recorded conversation and the actual appearance of the child on the stand for cross-examination, despite the difficulties in such cross-examinations. Clearly, a mechanical recording of the statement is better for the jury determining the truth than the necessarily selective testimony of a likely biased human observer, and since the law may be heading toward entirely exempting these interviews in most situations from Confrontation Clause scrutiny, recording may be the only realistic protection remaining.
considered above where the statement is elicited by a publicly disclosed investigative agent with the purpose of establishing facts about past events potentially relevant to later criminal prosecution. By contrast, coconspirator statements are apparently entirely exempt from scrutiny under the Confrontation Clause, whether made to true conspirators or to those who are government agents if their status as agents are unknown. Statements by children to non-law enforcement professionals who question the child with both investigative and non-law-enforcement interests in mind—what I term mixed purpose interviews—are in an uncertain category.

Moreover, Professor Berger’s approach of skepticism toward government generated hearsay and enabling the jury to make a better decision by the required creation and disclosure of mechanical recordings of the interviews remains sound even if not part of the testimonial statement approach of Sixth Amendment confrontation right. As the right matures through further rulings by the Court, perhaps a flexibility and nuance will be developed to supplement the right of confrontation, and ancillary protections of the type she suggests may be embraced.

If not, the fundamental idea is no less sound. It simply must get its support from another source and find its command in legislation or a different constitutional right, such as due process. My recent interests have prompted me to examine the role of the innocence movement in motivating and shaping criminal procedure reforms. One area of overlap that relates

40 In Giles v. California, 128 S. Ct. 2678, 2691 n.6 (2008), the Court stated that coconspirators statements “would probably never be . . . testimonial” because they must be made “in furtherance of the conspiracy.” This statement suggests an exclusive focus on the purpose or intention of the declarant in making the statement and suggests no different treatment when it is a government agent who elicits the coconspirator’s statement.
41 See Mosteller, supra note 6, at 968-75 (recognizing the uneven treatment of statements, including videotaped statements, made for multiple purposes and the capacity of those formulating the process to give the questioning an apparently non-investigate primary purpose); Robert P. Mosteller, Giles v. California: Avoiding Serious Damage to Crawford’s Limited Revolution, 13 LEWIS & CLARK L. REV. 675, 682 (2009) (noting an apparent or likely trend toward courts finding mixed purpose statements nontestimonial).
to Professor Berger’s insight is the control of informants who have played a role in the conviction of a number of innocent individuals for crimes they did not commit. My particular focus is on the danger of false informant testimony to those who are innocent of the specific crime charged but who are not strangers to crime."

The evidence that I find critical and subject to remedy is the testimony of informants who change or refine their version of events after contact with the police. Like the hearsay that is of concern to Professor Berger under the Confrontation Clause, this testimony sometimes bears the imprint of governmental agents who “turn” a criminal suspect into a cooperating witness by eliciting statements that incriminate the ultimate target of the prosecution. One of my remedies is based on an argument that the constitutional right under the Due Process Clause to have exculpatory information, including impeaching information, should have practical protection. Like Professor Berger, I argue that the police-citizen encounter should be recorded. “Because of their exposure to punishment and their strong desire to please the police and the prosecution, many of these informants are arguably vulnerable witnesses under her terminology."

There is no “magic bullet” to cure the dangers of informant testimony, particularly when defendants with past criminal involvement are concerned. Many of those defendants are clearly guilty and many of them are also threats to the insiders who might be able to offer incriminating testimony. Moreover, because of their general involvement in crime, investigative authorities will be receptive to the story the informants are offering, and informants will often have the raw material to fashion convincing false testimony from real or imagined past activities of the target or from the conduct of those who actually were responsible for the crime. Informants may be important to the justice system, but their testimony, which is sometimes false, is also dangerous to it.


43 See Mosteller, Producing Informant’s “First Drafts”, supra note 42, at 522.
44 Id.
45 Professor Bennett Gershman explicitly categorizes cooperating witnesses as vulnerable to suggestive questions along with children and eyewitnesses. See Bennett L. Gershman, Witness Coaching by Prosecutors, 23 CARDOZO L. REV. 829, 844 (2002).
Likely the only realistic protection for those who are innocent that is not too costly from a law enforcement perspective is to provide a more complete picture of the process to the jury. My proposal is, with Professor Berger, to require the recording of what I term the “first drafts” of informant testimony when such statements are produced after contact with criminal investigators.\(^{46}\)

If significantly inconsistent with the informant’s testimony, the production of those statements is already constitutionally required by the Brady doctrine.\(^{47}\) If consistent with the informant’s testimony, production should only benefit the government. Thus, there might seem little theoretical reason why the statements are not produced. However, practicality, adversarial incentives, and the fear that the process of “turning” the informant cannot withstand disclosure stand in the way.

Critically, most early statements are not recorded, and the fact that they are inconsistent remains unknown or undisclosed to the defense. It is the defense that is motivated to carefully examine the process and point out the changes in the story. Those on the police and prosecution side in many marginal situations do not have the mindset to notice and disclose inconvenient facts, which may be assumed to be innocuous under the sincerely held view that the defendant is guilty. More recording and disclosure should have a rightful role.

III. CONCLUSION: A CAREER CHARACTERIZED BY REASONED JUDGMENT

I close by observing that Professor Berger’s ideas are not only creative, but also have impact because wrapped into them are the intensity of her serious consideration and her attention to her craft as a practitioner, critic, and life-long student of the

\(^{46}\) So as not to inhibit police-suspect conversations and to provide a recognizable “trigger point,” I would impose the requirement of recording at the point the issue of a benefit to the informant is broached. See Mosteller, Producing Informant’s “First Drafts,” supra note 42, at 568-69.

\(^{47}\) See Brady v. Maryland, 373 U.S. 83 (1963) (finding the failure of the prosecution to provide potentially exculpatory evidence to the defense to be a due process violation). In United States v. Bagley, 473 U.S. 667 (1985), the Supreme Court characterized its ruling in United States v. Giglio, 405 U.S. 150 (1972), to be that “[i]mpeachment evidence . . . as well as exculpatory evidence, falls within the Brady rule,” and rejected a distinction between exculpatory evidence and impeachment evidence. See Bagley, 473 U.S. at 676.
law. For that set of characteristics I can find no better term, nor higher praise, than reasoned judgment." It is present in all that Professor Berger has done." It is a characteristic that is at the heart of what I believe all of us who devoted our careers to the law aspire.

In her long and exceptional career, Professor Margaret Berger has contributed so much to the law, and particularly to the development of the law of evidence, that any effort to mark her many contributions will be inadequate.\(^5\) I have tried to illustrate these contributions through her fundamental insights that we must rely on the imperfect institution of the jury to sort through our most difficult problems of proof, and for the jury to have a chance to do its task properly, it needs detailed and accurate information.

When the government is involved in creating evidence, particularly hearsay, the dangers of abuse are substantial. Declarants should be required to testify and be subject to cross-examination. In addition, one of our best and most realistic remedies for that abuse is to require that modern technology be employed to record in a verbatim fashion the transactions involved in that creation. By presenting that information, we

\(^{48}\) The writings of the constitutional scholar and my former colleague Jeff Powell inspire this accolade. In examining constitutional interpretation, he has coined the termed "constitutional virtues," which consist of (good) faith, integrity, humility, and candor. H. Jefferson Powell, Constitutional Virtues, 9 GREEN BAG 379, 389 (2006). In his recent book, Constitutional Conscience, Jeff expands on these fundamental concepts. See H. JEFFERSON POWELL, CONSTITUTIONAL CONSCIENCE: THE MORAL DIMENSION OF JUDICIAL Decision 100-01 (2008) (expanding the discussion and adding the virtue of acquiescence). A central example of what he believes should be the goal of everyone who attempts constitutional interpretation is the Attorney General in President Ulysses S. Grant's administration who was asked to render an opinion on a constitutional issue, which Powell believes was rendered in full adherence to his craft as a lawyer and his public duties. See id. ch. 3.

In a similar vein, I mean the characterization of "reasoned judgment" as truly a high compliment. As one who cares deeply about the practice of law and the intellectual exploration of ideas, Professor Berger has both mastered and been true to her craft.

\(^{49}\) Even before I entered law teaching, I held Professor Berger in high regard, first encountering her work in the masterful treatise on the Federal Rules of Evidence that she co-authored with Judge Jack Weinstein. See JACK B. WEINSTEIN & MARGARET A. BERGER, WEINSTEIN'S EVIDENCE (1975). I found that in almost all situations the treatise set forth with great precision, even in the early days of the operation of the rules, what the law was. Occasionally, it would deviate and state what the authors thought the law should be. These deviations were always creative, and they contained a remarkable measure of reasoned judgment. The courts may not have universally adopted their suggestions, but when they did not, it was most often a matter of judicial misjudgment.

\(^{50}\) For example, I have not even referred to her exceptional contribution to the analysis of the admission of scientific evidence and her service on multiple committees of the National Academy of Sciences.
better equip the jury to assess the value of the evidence and the value of the words of sometimes absent witnesses. Whether formally part of a confrontation right or recognized and guaranteed through other legal mechanisms, the insight is sound and manageable. It clearly shows the reasoned judgment of an insightful and committed legal scholar, who practiced her craft in ways that all in the field of evidence would hope to emulate.
Honoring Margaret Berger
with a Sensible Idea

INSISTING THAT JUDGES EMPLOY A BALANCING TEST BEFORE ADMITTING THE ACCUSED’S CONVICTIONS UNDER FEDERAL RULE OF EVIDENCE 609(a)(2)

Aviva Orenstein

INTRODUCTION

Impeachment of witnesses, though potentially edifying for the jury, can create a host of problems that undermine the fairness or accuracy of a trial. Certainly, the finder-of-fact (judge or jury) must be made aware of potential deficits in a witness’s credibility. The witness with poor eyesight who reports what she saw, the witness who has made a deal with the prosecutor, the witness with a reputation as a liar – all need to be impeached to reveal possible problems with their testimony. If the impeachment causes the jury to distrust the witness more than warranted, however, or, worse, to dislike the witness, valuable information may be inappropriately discounted. This problem is most acute when the witness in

---

1 Professor of Law, Indiana University Maurer School of Law. I want to express my deep regard for Margaret Berger as a generous colleague, wonderful teacher, and role model. I would like to thank Robin Ballard, Joshua Fix, Martha Marion, and Judy Reckelhoff for excellent research help. Thanks also to Seth Lahn, Hannah Buxbaum, Leandra Lederman, Sylvia Orenstein, William Popkin, Michael Risinger, Ted Sampsell-Jones, Eileen Scallen, and David Szonyi for their insights, even though and especially because they do not all agree with my conclusions. All mistakes are my own.

1 I focus on jurors because they tend not to be repeat players in the justice system and hence will have more difficulty with counterintuitive evidence principles. Jurors are more susceptible to certain types of visceral reactions and unfair prejudice than trained and experienced lawyers and judges, which is not to say that concerns about bias and irrationality are limited to juries alone. Finally, because judges have to rule on the admissibility of the evidence, the practical effect of exclusion (asking them to forget inadmissible evidence) is more questionable than in the case of jurors who are deprived of the evidence entirely.
question is also the accused in a criminal case. Impeachment that causes the jury to think of the accused as a bad person or someone who should still be in jail for prior crimes undermines the basic fairness of the trial.

This essay addresses a controversial form of impeachment: the use of prior convictions to impeach a witness’s character for truthfulness. It focuses on issues raised when the witness is the accused because in that instance, the rule is most interesting and its consequences most troubling. Under Federal Rule of Evidence 609(a), which was amended in 2006, any witness in a criminal or civil case may be impeached by her criminal convictions. Rule 609 divides the type of prior crimes with which a witness may be impeached into two categories: (1) felonies, and (2) any crimes involving dishonesty or false statement. Rule 609(a)(1) clearly provides a special

---

2 The “risk of unfair prejudice to a party in the use of [convictions] to impeach the ordinary witness is so minimal as scarcely to be a subject of comment.” Proposed FED. R. EVID. 609 advisory committee’s note, 51 F.R.D. 315, 392 (1971).
3 FED. R. EVID. 609(a).
4 The following is the text of Rule 609 reflecting the 2006 amendments. Crossed out material represents deletions; underlined material represents new language.

609. Impeachment by Evidence of Conviction of Crime
(a) General rule.
For the purpose of attacking the credibility character for truthfulness of a witness,
(1) evidence that a witness other than an accused has been convicted of a crime shall be admitted, subject to Rule 403, if the crime was punishable by death or imprisonment in excess of one year under the law under which the witness was convicted, and evidence that an accused has been convicted of such a crime shall be admitted if the court determines that the probative value of admitting this evidence outweighs its prejudicial effect to the accused; and (2) evidence that any witness has been convicted of a crime shall be admitted if it involved dishonesty or false statement, regardless of the punishment, if it readily can be determined that establishing the elements of the crime required proof or admission of an act of dishonesty or false statement by the witness.

FED. R. EVID. 609(a).

Aspects of Rule 609(a) reflect major changes in the legal landscape over the past three centuries. Until the late nineteenth century, people accused of crimes were not permitted to testify at all under the party-witness rule. Our rule against self-incriminating is actually half of a larger rule that the accused in a criminal matter was not allowed to testify either to incriminate or exculpate herself. Any rule for impeaching witnesses, therefore, did not apply to the accused, who was barred from taking the witness stand.

In a similar vein, historically, felons did not present impeachment conundrums; under the old English law, felonies were punishable by death, so all but those receiving pardons were unavailable to testify. Once some felonies became non-capital cases, the law had to decide what to do with convicted felons on the witness
balancing test that the judge must apply before admitting felonies to impeach a witness, with more stringent screening if the witness is the accused. This essay proposes that, even for convictions that involve dishonesty and false statement, the judge must screen for unfair prejudice before allowing such prior crimes to impeach the accused. In making this argument, this essay stakes a position that opposes the current interpretation of Rule 609 by federal courts, most state courts, and academia. It presents not merely a policy critique of Rule 609(a)(2), which many might agree with, but advocates that trial courts adopt a new interpretation and more limited application of this rule. In addition, perhaps this essay will persuade state legislatures to incorporate an explicit balancing test into state versions of their Rules of Evidence.

I. IMPEACHMENT OF THE ACCUSED WITH A PRIOR FELONY CONVICTION

Understanding how Rule 609 treats impeachment with crimes involving dishonesty or false statement requires that one first look at how the Rule treats felony convictions. Therefore, this essay first examines Rule 609(a)(1), which admits evidence of a felony conviction by the accused only if the probative value of the impeachment outweighs the prejudicial effect to the accused. Although this sounds very similar to Federal Rule of Evidence 403, which also employs language of probative value and prejudice, the balancing test for the accused imbedded in Rule 609(a) accomplishes something different. Rule 403 is a balancing test applied by the judge as a limited rule of exclusion, favoring admission of evidence; Rule 609, by contrast, is more restrictive. Further, under Rule 609(a)(1), the burden is on the prosecution to prove that such stand. The trend in evidence has moved from treating felons as entirely incompetent to testify to allowing them to testify with the impediment of disclosure of their prior crimes. See Green v. Bock Laundry Mach. Co., 490 U.S. 504, 511-12 (1989) (“As the law evolved, the absolute bar gradually was replaced by a rule that allowed such witnesses to testify in both civil and criminal cases, but also to be impeached by evidence of a prior felony conviction or a crimen falsi misdemeanor conviction.”).

5 Fed. R. Evid. 609(a)(1).

6 Fed. R. Evid. 403 provides: “Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.”
impeachment is more probative than prejudicial to the accused.\(^7\)

How exactly does a prior felony conviction impeach a witness? Ostensibly, by hearing about a fairly recent prior felony, the jury learns something about the character for truthfulness of a witness.\(^8\) The theory is that someone who would flout social norms by committing felonies might also be more likely to lie; the same anti-social tendency that led to felonious conduct could lead to perjury.\(^9\) At the best of times, such an implication is weak.\(^10\)

At the same time, information about the accused's previous felony has tremendous potential to make the trial unfair to the accused who takes the stand. The jurors will likely overvalue that information or otherwise misuse it. The mischief caused by evidence of the accused's convictions transcends the mere imputation of criminality.

The harm is magnified if the prior crime being used to impeach the testifying accused and the actual crime charged are similar. The jury may jump to the wrong type of propensity inference. Drawing an example from Rule 609(a)(1), imagine that someone is charged with armed robbery and takes the stand to deny her participation in the crime.

---

\(^7\) See Fed. R. Evid. 609 advisory committee's note to 1990 amendments ("Although the rule does not forbid all use of convictions to impeach a defendant, it requires that the government show that the probative value of convictions as impeachment evidence outweighs their prejudicial effect."). Also, all of Rule 609(a) applies only to crimes that occurred less than ten years from the "date of the conviction or of the release of the witness from the confinement imposed for that conviction, whichever is the later date." Fed. R. Evid. 609(b).

\(^8\) Rule 609 used to refer to credibility but was changed in 2006 to refer to "character for truthfulness." See Fed. R. Evid. 609 advisory committee's note to 2006 amendments.

\(^9\) This theory was articulated by Oliver Wendell Holmes, who as a Justice on the Supreme Judicial Court of Massachusetts, wrote in a civil case:

> [W]hen it is proved that a witness has been convicted of a crime, the only ground for disbelieving him which such proof affords is the general readiness to do evil which the conviction may be supposed to show. It is from that general disposition alone that the jury is asked to infer a readiness to lie in the particular case, and thence that he has lied in fact. The evidence has no tendency to prove that he was mistaken, but only that he has perjured himself, and it reaches that conclusion solely through the general proposition that he is of bad character and unworthy of credit.


\(^10\) See generally Richard D. Friedman, Character Impeachment Evidence: Psycho-Bayesian Analysis and a Proposed Overhaul, 38 UCLA L. REV. 637 (1991) (arguing that impeachment for character for truthfulness or prior convictions should never be allowed against a criminal defendant and explaining the limited probative value of such impeachment where the accused is the witness).
felony conviction, so the appropriate, if highly attenuated, implication from her prior conviction is that it casts light on her character for truthfulness. But what if the prior felony is for armed robbery? Even (or especially) with a limiting instruction, the jury is likely to gravitate to an impermissible inference—that the accused tends to commit armed robbery. This latter type of propensity evidence (as opposed to character for truthfulness) is clearly banned by the Evidence Rules. It is hard, however, to imagine how a jury could decline the unspoken invitation to think of the accused as a violent recidivist. Even if the evidence about the crime charged were weak, the jurors may unconsciously soften the burden of proof; they may be less scrupulous in weighing the evidence because the accused deserves punishment for past wrongs. In addition, they may wonder why the accused is at liberty if she has in the past committed armed assaults, believing that the accused is deserving of further punishment and preventive detention.

Given the low probative value of prior felony evidence, the extreme prejudice to the accused, and the fact that 609(a)(1) includes a special balancing test whereby felony convictions of the accused will be admissible only if the probative value of the prior felony outweighs the prejudicial effect to the accused, one might suppose that such evidence is rarely admitted. In fact, however, admission under Rule 609(a)(1) is a major factor in criminal trials, and troubling evidence exists that many accuseds do not take the stand primarily to avoid triggering this form of impeachment. In answering the common-sense (but devoid of the presumption of

11 See infra note 50 and accompanying text.
12 Rule 404(a) expressly excludes “evidence of a person’s character or trait of character” in order to prove “action in conformity therewith on a particular occasion.” FED. R. EVID. 404(a). Relatedly, Rule 404(b) provides: “Evidence of other crimes, wrongs, or acts is not admissible to prove the character of a person in order to show action in conformity therewith.” FED. R. EVID. 404(b).
13 This might very well be so where the police may have first come to suspect the accused because of her priors, and not because of any strong evidence linking her directly to the crime charged.
14 Witnesses other than the accused will have their felony convictions subjected to the more ubiquitous and permissive Rule 403 balancing test.
15 See John H. Blume, The Dilemma of the Criminal Defendant with a Prior Record—Lessons from the Wrongfully Convicted, 5 J. EMPIRICAL LEGAL STUD. 477, 491 (2008) (documenting that of the factually innocent accused in his data set who failed to take the witness stand, 91% had prior convictions and “[i]n almost all instances . . . counsel for the wrongfully convicted defendant indicated that avoiding impeachment was the principal reason the defendant did not take the stand”); see also Jeffery Bellin, Circumventing Congress: How The Federal Courts Opened The Door To Impeaching Criminal Defendants With Prior Convictions, 42 U.C. DAVIS L. REV. 289, 293 (2008).
innocence) question why an innocent accused would not take the witness stand in her own defense, one reason is that the accused is afraid of the effects of being impeached by a prior conviction.

As a functional matter, courts regularly misapply Rule 609(a)(1) by allowing prosecutors to impeach the accused with felonies in ways that conflict directly with the letter, spirit, and history of the Rule. In one subset of cases, however, admission-happy courts tend to exercise caution and restraint. Courts tend to take seriously the factor concerning the similarity between the past conviction and the charged crime, and exclude evidence of prior felonies to the extent that they resemble the crime charged. Courts recognize that the chances of jury confusion and unfair prejudice are particularly egregious in such cases. Under the dominant interpretation of Rule 609(a)(2), however, courts currently have no opportunity to consider this prejudice. The next Part presents Rule 609(a)(2) and sets the stage for a discussion whether impeachment for crimes involving dishonesty or false statement under Rule 609(a)(2) is subject to Rule 403 balancing.

16 Recently, Jeffrey Bellin criticized the federal courts’ “routine admission of defendants’ prior convictions” under Rule 609(a)(1) as contravening congressional intent. Bellin, supra note 15, at 293. Such a framework is the intellectual descendant of the old, English common law tradition barring felons from testifying at all. See id. at 296-97 nn. 21, 23-24. Bellin outlined the “decidedly pro-impeachment,” five-factor analytical framework that “places an almost insurmountable burden on defendants attempting to exclude prior convictions.” Id. at 293.

17 See 28 CHARLES A. WRIGHT & VICTOR J. GOLD, FEDERAL PRACTICE & PROCEDURE § 6134 (1993) (“The danger of prejudice is enhanced if the witness is the accused and the crime was similar to the crime now charged, since this increases the risk that the jury will draw an impermissible inference under Rule 404(a).”). See, e.g., United States v. Bagley, 772 F.2d 482, 488 (9th Cir. 1985) (prior conviction for bank robbery excluded as impeachment of accused in current bank robbery charge because “there is a substantial risk that all exculpatory evidence will be overwhelmed by a jury's fixation on the human tendency to draw a conclusion which is impermissible in law: because he did it before, he must have done it again”); United States v. Joe, 07 Cr. 734 (JFK), 2008 U.S. Dist. LEXIS 55036, at *11 (S.D.N.Y. July 21, 2008) (excluding prior conviction because “the prior conviction for firearms possession is nearly identical to the conduct charged in Count One, [thus] the jury may infer unfairly that Defendant has a propensity to commit firearms offenses”); United States v. Jaramillo, No. 1:05-CR-13, 2007 U.S. Dist. LEXIS 38016, at *7-8 (D. Utah May 24, 2007) (accused's prior convictions for possession of controlled substances would not be admissible because their probative value “would be outweighed by their prejudicial effect because the jury may consider them to be evidence that he committed the possession crimes charged rather than merely probative of his character for truthfulness”).
II. IMPEACHMENT FOR CRIMES INVOLVING DISHONESTY OR FALSE STATEMENT

Rule 609(a)(2), last amended in 2006, provides that “evidence that any witness has been convicted of a crime shall be admitted regardless of the punishment, if it readily can be determined that establishing the elements of the crime required proof or admission of an act of dishonesty or false statement by the witness.” Historically, these types of crimes, labeled by the common law as crimen falsi, were considered particularly probative for impeachment purposes; because the prior convictions actually have something to do with dishonesty, the inference that they reflect on the character for truthfulness of the accused seems reasonable.

Rule 609(a)(2), however, presents some interesting variations: it makes no distinction between felonies and misdemeanors; it does not differentiate the accused from other witnesses; and, most importantly for this essay, it makes no mention of any balancing test.

Although there is textual, historical, doctrinal, and policy evidence for the proposition that Rule 609(a)(2) allows for no balancing, this essay advocates that this orthodoxy be reconsidered. It presents the arguments favoring the position that Rule 609(a)(2) permits no balancing, and then offers countervailing reasons that demonstrate why balancing is not only permissible, but necessary.

18 Fed. R. Evid. 609(a)(2).
19 Scholars, rulemakers, and courts have disagreed about what types of crimes fall into this special category regarding dishonesty and false statement, with theft and receipt of stolen property being areas of contention. Compare United States v. Gunter, 551 F.3d 472, 483 (6th Cir. 2009) (appellate court refused to consider trial court ruling “that because theft is an offense involving dishonesty under Tennessee state law, the convictions could be used for impeachment purposes under Rule 609(a)(2)” where the defendant chose not to testify in light of the trial court’s ruling and therefore waived his right to appeal the ruling), and U.S. Xpress Enters. v. J.B. Hunt Transp., 320 F.3d 809, 816-17 (8th Cir. 2003) (trial court did not abuse discretion in finding that receiving stolen property was a crime involving dishonesty within the meaning of Rule 609(a)(2)), with United States v. Glenn, 667 F.2d 1269, 1273 (9th Cir. 1982) (“Generally, crimes of violence, theft crimes, and crimes of stealth do not involve ‘dishonesty or false statement’ within the meaning of rule 609(a)(2),”), and United States v. Foster, 227 F.3d 1096, 1100 (9th Cir. 2000) (“We therefore hold that receipt of stolen property is not per se a crime of dishonesty for purposes of Rule 609(a)(2), and conclude that the district court erred in treating it as such.”).

The 2006 amendment was intended to “give effect to the legislative intent to limit the convictions that are to be automatically admitted under subdivision (a)(2)” to such cases where the act of falsity or dishonesty is obvious from the nature of the crime charged. Fed. R. Evid. 609 advisory committee’s note to 2006 amendments.
III. TRADITIONAL ARGUMENTS IN FAVOR OF NO BALANCING

A. A Plain Reading Analysis Indicates No Rule 403 Balancing

The text of Rule 609(a)(2) currently states that convictions for crimes of dishonesty or false statement by the witness “shall” be admitted. They are not “admissible,” nor is the word “may” used, but instead the imperative “shall” is employed.

Nowhere does the text of Rule 609(a)(2) call for any balancing. Given the fact that other parts of Rule 609 include three different balancing tests, it seems unlikely that Congress’s failure to add a balancing test to Rule 609(a)(2) was a mere inadvertent omission.

Using another cannon of construction, one could argue that Rule 403 is a general rule that must give way to the more particular Rule 609. In United States v. Kiendra, the First Circuit explained:

20 [FED. R. EVID. 609(a)(2)].

21 “Shall” is inherently ambiguous—in that it can mean “must,” but also can mean “should” or “will.” For that reason, in its recent restyling of the Federal Rules of Civil Procedure, the word shall was excised entirely, and the proposed “restyled” Rule 609(a)(2) substitutes the word “must” for “shall.” See ADVISORY COMMITTEE ON EVIDENCE RULES, AGENDA FOR COMMITTEE MEETING 136 (Oct. 23-24, 2008), available at http://federalevidence.com/pdf/2009/01-Jan/RestyleFRE501-706.pdf. It is fair to say even before the proposed amendment that courts have read Rule’s 609(a)(2)’s “shall” as an imperative. The switch to “must” for restyling purposes, even if it does go through, is not fatal to the argument of this essay. The restyling of the rules cannot introduce substantive changes, so if the argument that balancing is permissible and should be encouraged is correct under the current rule, nothing should change with the restyling. The template Committee Note to each of the restyled rules reads: “The language of Rule [ ] has been amended as part of the restyling of the [ ] Rules to make them more easily understood and to make style and terminology consistent throughout the rules. These changes are intended to be stylistic only.” Memorandum from Robert L. Hinkle, Chair, Advisory Comm. on Evidence Rules, to Hon. Lee H. Rosenthal, Chair, Standing Comm. on Rules of Practice & Procedure 3 (May 12, 2008), available at http://federalevidence.com/pdf/2008/07-July/ECRpt%20May%202008-1.pdf. To qualify as a substantial change, a revision “changes the structure of a rule so as to alter the way in which courts and litigants have thought about, and argued about, questions of admissibility.” Id. at 2. The change to “must” confirms, however, that my proposed reading of Rule 609(a)(2) is unorthodox.

22 The three tests are: (1) the special balance for the accused in Rule 609(a)(1); (2) Rule 403 balance for all other witnesses in Rule 609(a)(1); and (3) the high hurdle for admission balance under 609(b) whereby the court may admit a stale conviction only if the court determines, “in the interests of justice, that the probative value of the conviction supported by specific facts and circumstances substantially outweighs its prejudicial effect.” FED. R. EVID. 609(b).

Rule 403 is a general provision intended to govern a wide landscape of evidentiary concerns; Rule 609 is a narrow provision intended to regulate the impeachment of witnesses who have been convicted of prior crimes . . . “Specific terms prevail over the general in the same or another statute which otherwise might be controlling.”24

Thus, it is argued that Rule 403 was not designed to override the more specific Rule 609; rather it was “designed as a guide for the handling of situations for which no specific rules have been formulated.”25

B. No Balance Is Necessary Because Crimes Under 609(a)(2) Are Particularly Probative

Historically, crimes involving false statement or dishonesty best represent the policy underlying impeachment for character for truthfulness.26 One might argue that balancing is least necessary for these types of crimes because the probative value of crimes involving false statement or dishonesty is very high.27 Unlike garden-variety felonies, which merely show the witness’s anti-social tendencies, convictions contemplated by 609(a)(2) demonstrate that the accused is willing to lie and deceive. Additionally, advocates of employing Rule 609(a)(2) without any balancing test note that the recent amendment further narrowed the scope of this Rule, limiting it to the types of convictions in which falsity and dishonesty would be readily apparent from the elements of the crime. Hence, it is argued, such cases are increasingly few and more tailored to the core concern of Rule 609—imputing character for untruthfulness.

---

25 Id. (quoting Fed. R. Evid. 403 advisory committee’s note); see also Palm Beach County Canvassing Bd. v. Harris, 772 So. 2d 1220 (Fla. 2000); Fed. R. Evid. 403 advisory committee’s note.
26 In the original debates over Rule 609, the United States House of Representatives advocated a version of the Rule that limited such impeachment to convictions involving proof or admission of an act of dishonesty or false statement. See Diggs v. Lyons, 741 F.2d 577, 580 (3d Cir. 1984) (reviewing the legislative history of Rule 609) (citing 120 Cong. Rec. 2381).
C. The Legislative History Indicates that the Drafters of Rule 609(a)(2) Never Anticipated Balancing

Rule 609 was hotly debated in Congress. An earlier draft of Rule 609 included subsection 609(a)(3), which would have allowed the court to exclude any type of conviction if the probative value was substantially outweighed by its prejudicial effect. This subsection was severely criticized and apparently rejected. Another school of thought advocated no balancing whatsoever for any prior crimes. Therefore, those who believe there is no balancing under Rule 609(a)(2) argue that the Rule represents a purposeful decision not to engage in such balancing.

Those who insist on applying Rule 609 without a balancing test point out that the Rule was “the product of extensive Congressional attention and considerable legislative compromise, clearly reflecting a decision that judges were to have no discretion to exclude crimen falsi.” After an extensive review of the legislative history, the Court in United States v. Wong concluded that Rule 609(a)(2) “unambiguously demonstrates that a judge has no authority to prohibit the government’s effort to impeach the credibility of a witness by questions concerning a prior crimen falsi conviction.”

This view is supported by the Conference Report on Rule 609, which explained:

The admission of prior convictions involving dishonesty and false statement is not within the discretion of the Court. Such convictions are peculiarly probative of credibility and, under this rule, are always to be admitted. Thus, judicial discretion granted with respect

---


29 117 CONG. REC. 29, 894-95 (1971). When the Supreme Court officially promulgated the Federal Rules of Evidence and transmitted them to the Congress, section 609(a)(3) had disappeared. See United States v. Wong, 703 F.2d. 65, 67 (3d Cir. 1983).

30 See Diggs, 741 F.2d at 579 (reviewing legislative history).

31 Id. at 581 (quoting United States v. Kiendra, 663 F.2d 349, 355 (1st Cir. 1981)); United States v. Wong, 703 F.2d 65 (3d Cir. 1983); United States v. Toney, 615 F.2d 277, 278, 280 (5th Cir. 1980) (“Congress thoroughly considered the pros and cons of the mandatory admissibility of limited types of prior crimes evidence and determined that in certain cases it was to be the rule. Rule 403 simply has no application where impeachment is sought through a crimen falsi.”).

32 703 F.2d 65, 68 (3d Cir. 1983).
to the admissibility of other prior convictions is not applicable to those involving dishonest or false statement.\textsuperscript{33}

\section*{D. Courts Are Unanimous in Ruling that Rule 609(a)(2) Allows for No Balancing}

In \textit{Green v. Bock Laundry Machine Co.}, the Supreme Court observed of Rule 609(a), in dicta: “With regard to subpart (2), which governs impeachment by \textit{crimen falsi} convictions, it is widely agreed that this imperative, coupled with the absence of any balancing language, bars exercise of judicial discretion pursuant to Rule 403.”\textsuperscript{34} All circuits have also come to this same conclusion,\textsuperscript{35} though they debate the parameters of the rule.\textsuperscript{36}

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{33} H.R. Rep. No. 1597, at 9 (1974), \textit{reprinted in} 1974 U.S.C.C.A.N. 7098, 7102. Similarly, the original advisory committee’s note explained that the “admission of prior convictions involving dishonesty and false statement is not within the discretion of the court. Such convictions are peculiarly probative of credibility and, under this rule, are always to be admitted.” FED. R. EVID. 609 advisory committee’s note.
\item \textsuperscript{34} 490 U.S. 504, 525-26 (1989), \textit{superseded on other grounds by} 1990 amendment to FED. R. EVID. 609(a).
\item \textsuperscript{35} See, e.g., SEC v. Sargent, 229 F.3d 68, 80 (1st Cir. 2000) (“[W]e have plainly held that district courts do not have discretion to exclude prior convictions involving dishonesty or false statements.”); United States v. Estrada, 430 F.3d 606, 615-16 (2d Cir. 2005) (“[C]rimes involving ‘dishonesty or false statement,’ whether felonies or misdemeanors, \textit{must} be admitted under Rule 609(a)(2) as being per se probative of credibility.”); Walker v. Horn, 385 F.3d 321, 333 (3d Cir. 2004) (“[I]f the prior conviction involved dishonesty or false statements, the conviction is automatically admissible insofar as the district court is without discretion to weigh the prejudicial effect of the proffered evidence against its probative value.”) (internal citation omitted); United States v. Kelly, 510 F.3d 433 (4th Cir. 2007), \textit{cert. denied} Kelly v. United States, 552 U.S. 1329 (2008) (“A trial judge has no discretion to exclude evidence that qualifies under this rule (609(a)(2)).”) (internal citation omitted); United States v. Harper, 527 F.3d 396 (5th Cir. 2008) (“Crimes qualifying for admission under Rule 609(a)(2) are not subject to Rule 403 balancing and must be admitted.”); United States v. Morrow, 977 F.2d 222, 228 (6th Cir. 1992) (en banc) (“Rule 609(a)(2) . . . clearly limits the discretion of the court by mandating the admission of crimes involving dishonesty or false statements.”); Kunz v. DeFelice, 538 F.3d 667, 675 (7th Cir. 2008) (“Rule 609(a)(2) . . . does not incorporate Rule 403.”); United States v. Collier, 527 F.3d 695 (8th Cir. 2008) (“Evidence of a conviction requiring proof or admission of an act of dishonesty or false statement is automatically admissible and not subject to Rule 403 balancing by the court.”) (internal citations omitted); United States v. Harris, 738 F.2d 1068, 1073 (9th Cir. 1984) (“[C]rimes involving dishonesty and fraud are automatically admissible for impeachment purposes under Fed. R. Evid. 609(a)(2).”) (internal citations omitted); United States v. Begay, 144 F.3d 1336, 1338 (10th Cir. 1998) (“Rule 403 balancing applies unless the prior crime involves dishonesty or false statements.”). Although I could find no current case in the Eleventh Circuit, a Fifth Circuit case before the circuit split would seem to apply. \textit{See} United States v. Williams, 642 F.2d 136, 140 (5th Cir. 1981) (“[B]ribery is a \textit{crimen falsi} in that it involves dishonesty. Hence, it is automatically admissible.”) (internal citation omitted).
\item \textsuperscript{36} There were certainly debates in the common law surrounding Rule 609(a)(2), but they concerned the scope of \textit{crimen falsi}. Historically, some courts extended Rule 609(a)(2) to crimes such as drug use or robbery, crimes in which the aspect of false statement or dishonesty was highly questionable. Even before the 2006
The treatise writers also agree.\textsuperscript{37} This unanimity was not always the case under the prior common law,\textsuperscript{38} and a number of courts considered the question open during the late 1970s.\textsuperscript{39} Before the codification of the evidence rules, the common law, while favoring the use of impeachment for crimes of dishonesty and false statement, did not deprive judges of discretion to disallow such evidence.\textsuperscript{40}

\begin{enumerate}
\item[] E. The Fact of Recent Amendment Indicates that Judicial Interpretations of Rule 609(a)(2) Were Acceptable
\end{enumerate}

The Federal Rules of Evidence are an odd hybrid of statutes passed by Congress, and rules and amendments promulgated via the Rules Enabling Act process.\textsuperscript{41} In 2006, Rule 609 was amended via the rulemaking process, and included important changes to Rule 609(a)(2), yet nothing was altered concerning the apparent absence of balancing. Given the unanimous judicial authority indicating no balancing for Rule 609(a)(2), this would seem like an endorsement of those judicial amendment, the modern trend had been to limit the types of crimes admissible under this rule, precisely because there was perceived to be no opportunity for balancing. See, e.g., United States v. Foster, 227 F.3d 1096, 1100 (9th Cir. 2000) (holding that receipt of stolen property is not per se a crime of dishonesty); cf. Cree v. Hatcher, 969 F.2d 34, 37 (3d Cir. 1992) (“Because the district court lacks the discretion to engage in balancing, 609(a)(2) must be interpreted narrowly to apply to only those crimes that, in the words of the Conference Committee, bear on a witness’s propensity to testify truthfully.”). The 2006 amendment limits the types of crimes even further. Courts also debate the level of detail that could be presented about the conviction. United States v. Sine, 493 F.3d 1021, 1036 n.14 (9th Cir. 2007) (noting that convictions admitted for impeachment may not include collateral details). Cf. Commerce Funding v. Comprehensive Habilitation Servs., 01 Civ. 3796, 2005 U.S. Dist. LEXIS 7902, at *26-27 (May 2, 2005) (in a civil case, admitting prior crime under 609(a)(2) but limiting the underlying facts using Rule 403).

\textsuperscript{37} See WRIGHT & GOLD, supra note 17, at § 6135 (“[S]ubdivision (a)(2) neither requires nor permits balancing under Rule 403 or any other test.”); see also infra note 78.

\textsuperscript{38} See United States v. Toney, 615 F.2d 277 (5th Cir. 1980) (Tuttle, J., dissenting); Gordon v. United States, 383 F.2d 936, 940-41 (D.C. Cir. 1967) (allowing balancing for all prior convictions of the accused).


\textsuperscript{40} See, e.g., Gordon, 383 F.2d at 939-40 (setting out criteria for the admission of prior crimes, but applying no absolute rules concerning crimes of dishonest or false statement).

\textsuperscript{41} Congress rejected the Evidence Rules as promulgated in 1975 and instead codified its own version as statutes. Other amendments have been made as statutory additions to the Evidence Rules. See, e.g., FED. R. EVID. 413-415.
interpretations, albeit by a body different from the original legislature that enacted Rule 609.\textsuperscript{42}

IV. WHY THERE MUST BE SOME BALANCING IN APPLYING RULE 609(a)(2)

Given the consensus among judges and academics supported by a plain reading of the text, dicta from the United States Supreme Court, and the Conference Report on Rule 609 that Rule 609(a)(2) does not allow balancing, what is the justification for the contrary view? The counter-arguments reflect concerns for basic fairness and the structural integrity of the administration of evidence rules. Although subjecting Rule 609(a)(2) to Rule 403 balancing is not the most natural interpretation of the plain meaning, it is a reasonable one that must be preferred to guarantee due process.

A. Without Balancing, There Is Potential for Intolerable Unfairness to the Accused

Although one could aptly criticize courts’ balancing for felony convictions under 609(a)(1), courts do tend to follow one protective principle: the exclusion of prior felony convictions where the accused is charged with a similar crime.\textsuperscript{43} By contrast, under the current interpretation of Rule 609(a)(2), no such consideration, no matter how extreme the potential prejudice, is entertained.

Where the crime charged and the prior conviction are the same, the prejudice is overwhelming. For instance, in United States v. Wilson,\textsuperscript{44} the accused was charged with conspiring to defraud the United States government of tax revenue and was impeached with his prior conviction for failure to file tax returns. No Rule 403 balancing test was

\textsuperscript{42} Cf. FDA v. Brown & Williamson Tobacco Corp., 529 U.S. 120, 156 (2000) (“Congress’ tobacco-specific legislation has effectively ratified the FDA’s previous position that it lacks jurisdiction to regulate tobacco.”).

\textsuperscript{43} See supra note 17 and accompanying text; see also 4 JACK B. WEINSTEIN & MARGARET A. BERGER, WEINSTEIN’S FEDERAL EVIDENCE § 609.05[3][d] (Joseph M. McLaughlin ed., Matthew Bender 2d ed. 1997) (“When a prior crime committed by the accused is similar to the one with which the accused is charged, the prejudicial effect of a prior conviction admitted for impeachment may well outweigh its probative value. Consequently, prior convictions for the same or similar crimes are admitted sparingly.”).

\textsuperscript{44} 985 F.2d 348, 350-51 (7th Cir. 1993).
allowed or conducted. The permitted chain of inference, that since the accused evaded taxes before, he is probably a liar who is not telling the truth about evading taxes now, is preposterously obscure and ultimately toxic. Realistically, such impeachment invited the jury to think of the accused as a recidivistic tax dodger. Other cases, such as a bank fraud case where the accused’s prior conviction for bank larceny was admitted, mail fraud where the accused’s prior mail fraud was admitted, and counterfeiting where the accused’s prior conviction for counterfeiting was admitted, similarly fit this template. In none of these, however, did the court apply a Rule 403 balancing test.

The potential for extreme unfairness stems from two factors: (1) the effect on the jury; and (2) the consequent strategic decision, which many of the accuseds will make, not to testify at all. The presumption of innocence is a hard principle to effectuate even under the best of circumstances. The jurors already will have a tough time remembering that the person sitting at the defense table, whom the police have arrested and the government is prosecuting, must be presumed innocent until proven otherwise. This presumption becomes much more difficult when the jurors learn that the accused has a criminal past and, in fact, a record for doing the exact same thing. The jurors will label the accused as a criminal with a specialty crime, and not limit their skepticism to just the question of the accused’s character for truthfulness.

45 Id. at 351 (citations omitted).
46 This is also the fact pattern in United States v. Tanaka, 204 Fed. Appx. 705, 706 (9th Cir. 2006). The accused was charged with convictions for structuring transactions to evade currency reporting requirements and willful failure to file a tax return. The court admitted his prior tax conviction under Rule 609(a)(2). The court explained: “Failure to file a tax return is a crime involving dishonesty or false statement, and crimes involving dishonesty are automatically admissible for impeachment purposes under Federal Rule of Evidence 609(a)(2), and no balancing of prejudice is required.” Id.
48 United States v. Kuecker, 740 F.2d 496, 498, 502 (7th Cir. 1984); United States v. Toney, 615 F.2d 277, 278, 280 (5th Cir. 1980) (“Congress meant what it said in rule 609(a)(2) that the fact of a prior conviction for an offense such as mail fraud is always admissible for impeachment purposes.”).
50 Limiting instructions will not solve the problem. A limiting instruction would theoretically focus the jury on the appropriate inference of character for truthfulness. The following is typical: “Th[e] [defendant’s] earlier conviction was brought to your attention only as one way of helping you decide how believable his
Furthermore, the jury may soften the standard of proof. Once jurors hear that the accused has been convicted of forgery, mail fraud, or tax evasion in the past, the jurors’ concern about making a mistake declines significantly. Even if the accused did not commit the crime this time, it is not as if she is without blame for some similar activity in the past, so an erroneous conviction would not be a tragedy of an innocent person falsely convicted. Finally, if the prior conviction reflects conduct more heinous than the charged crime, the jury could be distracted or seek to punish the accused further.

This line of thinking is thoroughly predictable; therefore, the accused must consider the nature of the impeachment she will experience if she takes the stand in her own defense. If the judge has no power to exclude the accused’s prior crime involving dishonesty or false statement, no matter how prejudicial, the accused may simply decline to exercise her right to testify and avoid impeachment altogether. Thus, the accused is presented with an impossible dilemma. If she testifies, the jury will form negative assessments of her criminality generally and her propensity to commit certain crimes. On the other hand, failure to take the stand also presents a huge, unfair burden on the accused because jurors tend to believe that criminal defendants who do not testify are more likely to be guilty.\textsuperscript{51} 

\textsuperscript{51} Testimony was. You cannot use it for any other purpose. It is not evidence that he is guilty of the crime that he is on trial for now.” O’Malley, Greig & Lee, 1A Federal Jury Practice & Instructions § 15.08 (5th ed. 2007) (listing this instruction from Sixth Circuit and providing other examples by Circuit) (cited in Bellin, supra note 15, at 300 n.37). Scholars have long noted the ineffectiveness of limiting instructions; in fact, such instructions tend to draw attention to the evidence and may inadvertently increase the unfairness of its use. Joel Lieberman, Understanding the Limits of Limiting Instructions: Social Psychological Explanations for the Failures to Instructions to Disregard Pretrial Publicity and Other Inadmissible Evidence, 6 Psychol. Pub. Pol’y. & L., 677, 677-86, (2000) (empirical research has demonstrated that judicial admonitions to ignore evidence are relatively ineffective and sometimes produce a backfire effect, resulting in jurors being more likely to rely on inadmissible information after they have been specifically instructed to disregard it); J. Alexander Tanford, Thinking About Elephants: Admonitions, Empirical Research and Legal Policy, 60 UMKC L. Rev. 645, 651 (1992) (discussing the psychological futility of limiting instructions). 

See David Shaffer & Thomas Case, On the Decision to Testify in One’s Own Behalf: Effects of Withheld Evidence, Defendant’s Sexual Preferences, and Juror Dogmatism on Juridic Decisions, 42 J. Personality & Soc. Psychol. 335 (1982) (indicating that defendants who invoked the Fifth Amendment (either on the stand or by declining to take the stand) were judged more likely to be guilty and more deserving of conviction than their counterparts who took the stand and answered all questions. “It appears as if many of our subject-jurors chose to disregard that judge’s instructions and act on an impression that an innocent person who had nothing to hide would surely not resort to such legal chicanery as a Fifth Amendment plea.”).
Although the accused often faces the dilemma of whether to testify and be impeached with a former crime or to forgo the opportunity to testify altogether, that decision, when it involves felonies under 609(a)(1), is made in light of a judge’s belief that the prior conviction was more probative than prejudicial. Without introducing Rule 403 as a judicial screening device, prior crimes that are highly prejudicial would be admitted with no such check. This is a loss not only for the accused, but also for the whole justice system because the criminal defendant likely has valuable information that will increase trial accuracy.\footnote{See Ted Sampsell-Jones, Making Defendants Speak, 93 MINN. L. REV. 1327, 1328-29 (2009) (proposing changes in evidence law to encourage the accused to take the witness stand).}

**B. The Probative Value of Any Prior Conviction Is Diminished in the Case of the Accused**

Assuming, arguendo, that prior convictions for dishonesty or false statement are generally probative of a witness’s character for truthfulness, the probative value of such impeachment declines significantly when the witness is the accused.\footnote{Theoretically, it is possible to believe an accused’s denials slightly less when one hears that the accused who denied moving survey markers on federal land committed perjury in the past, United States v. Caudle, 48 F.3d 433, 435 (9th Cir. 1995), or the accused who is on trial for bank fraud was convicted of tampering with an odometer. United States v. Harris, 512 F. Supp. 1174, 1175-78 (D. Conn. 1981).} Jurors are probably already skeptical of the testimony of a defendant who claims that she did not commit the charged crime. It is natural to suppose that the accused might lie about her conduct to avoid punishment.\footnote{See Friedman, supra note 10 at 638, 659 (arguing that prior bad acts of the accused are “almost certain to yield no significant new information about his truhtelling inclination in the specific case” because “a rational jury usually will conclude, even without character impeachment, that the accused has a strong interest in lying and little compunction against doing so. Character impeachment evidence is overkill”); Gene R. Nichol, Jr., Prior Crime Impeachment of Criminal Defendants: A Constitutional Analysis of Rule 609, 82 W. VA. L. REV. 391, 408 (1980) (“Greater incentive to deceive can hardly be imagined [than a defendant’s interest in acquittal] and this motive and propensity are well understood and recognized by each member of the jury.”) (quoted in Jeffrey Bellin, Circumventing Congress: How the Federal Courts Opened the Door to Impeaching Criminal Defendants with Prior Convictions, 42 U.C. DAVIS L. REV. 289, 289 n.32 (2008)); Sampsell-Jones, Making Defendants Speak, supra note 52, at 1367 (“[I]n every criminal case, a defendant’s interest in the outcome is obvious, therefore lessening the need for other impeachment evidence.”) (citing 1 MCCORMICK ON EVIDENCE § 39, at 173 (Kenneth S. Brown ed., 6th ed. 2006)).} Therefore, the added information that the accused actually has committed a prior crime of dishonesty or false statement adds significant...
prejudice without adding any additional probative value or promoting a more informed verdict.\textsuperscript{55}

C. \textit{From a Linguistic Perspective, Rule 609(a)(2) Could Be Read as Subject to Rule 403}

Although Rule 609(a)(2) says that crimes of falsity and dishonesty “shall be admitted,” it is nevertheless still possible to read that form of impeachment as being subject to the pervasive balance of Rule 403. Even with the transition to “must” in the restyled rules,\textsuperscript{56} one could argue that it must be admitted subject to Rule 403. To read Rule 609(a)(2) as outside the reach of Rule 403 would mean that it is an evidence rule in a category all its own, the only one entirely insulated from judicial intervention on basic fairness grounds. Given that unusual status, the presumption should be that all rules are subject to Rule 403 unless the balance is affirmatively rejected or another balance is proposed in its place. “Although one must be somewhat of an interpretative funambulist to walk between the conflicting demands of these Rules in order to arrive at a resolution,”\textsuperscript{57} this is a credible, if strained reading that will effectuate vital evidence policy and address the due process concerns raised below.\textsuperscript{58}

\textsuperscript{55} Some states that otherwise have adopted the Federal Rules of Evidence as their own deviate from the template of Rule 609. Examples include: Arizona, Ariz. R. Evid. 609(a)(1) (which applies the same test to felonies and crimes involving dishonesty or false statement); Tennessee, Tenn. R. Evid. 609(a)(2) (“If the witness to be impeached is the accused in a criminal prosecution . . . the court upon request must determine that the conviction’s probative value on credibility outweighs its unfair prejudicial effect on the substantive issues.”); and Hawaii, Haw. R. Evid. 609(a) (limiting all impeachment to convictions involving dishonesty and false statements, but only allowing such impeachment against the accused when the accused has first presented evidence supporting her credibility). Montana allows no impeachment by conviction at all. Mont. R. Evid. 609.

\textsuperscript{56} See supra note 21.


\textsuperscript{58} The Supreme Court of Iowa has taken this policy a step further, and requires more stringent 609(a)(1) balancing of 609(a)(2) convictions, despite the fact that Iowa’s rules of evidence read identically to the pre-2006 federal rules for the purposes of Rule 609(a). See State v. Axiotis, 569 N.W.2d 813, 815 (Iowa 1997) (“Iowa rule of evidence 609, subsection 609(a)(2) provides concerning impeachment that ‘evidence that any witness has been convicted of a crime shall be admitted if it involved dishonesty or false statement . . . .' The trial court also must determine ‘that the probative value of admitting this evidence outweighs its prejudicial effect to the accused.’”).
D. There is Precedent in the Evidence Rules for Acknowledging the Potential Due Process Problems with Similar Act Character Evidence of the Accused and for Applying Rule 403 Despite the Absolutist Language

A more recent debate about the applicability of Rule 403 to another set of Evidence Rules offers insight into how Rule 609(a)(2) should be interpreted. Federal Rules of Evidence 413-414, enacted by statute in 1995, have been subject to Rule 403 balancing even though Rule 403 is not mentioned in those Rules. The language of Rule 413 commands that “[i]n a criminal case in which the defendant is accused of an offense of sexual assault, evidence of the defendant’s commission of another offense or offenses of sexual assault is admissible, and may be considered for its bearing on any matter to which it is relevant.” Arguably, the language “is admissible” represents an unmodified command to admit the evidence, prohibiting judges from exercising their discretion to balance between probative value and unfair prejudice. However, every court that has considered the question has held that Rule 403 applies.\(^\text{59}\)

Admittedly, Rules 413-414 do not present a perfect analogy. The phrase “is admissible” is less absolute than the command of Rule 609(a)(2) “shall be admitted.”\(^\text{60}\) At least part of the legislative history of Rules 413-414 anticipated the use of Rule 403 balancing, even though Rule 403 itself is not mentioned in the text of Rules 413-414.\(^\text{61}\) And, no variation of

\(^\text{59}\) FED. R. EVID. 413. The language of Rule 414 is identical, except that Rule 413 is for sex; Rule 414 is for child molestation. FED. R. EVID. 414. The restyled rules, perhaps in light of the jurisprudence surrounding Rules 413-414 read: the court may admit evidence that the defendant committed any other sexual assault.

\(^\text{60}\) See Aviva Orenstein, Deviance, Due Process, and the False Promise of Federal Rule of Evidence 403, 90 CORNELL L. REV. 1487, 1517-19 (2005) (noting the courts’ holdings that Rule 403 applies but expressing concern that courts do no apply the balancing test in a meaningful fashion, but instead engage in Rule “403-lite”).

\(^\text{61}\) See United States v. Sumner, 119 F.3d 658, 661-62 (8th Cir. 1997).

\(^\text{62}\) Then representative Kyl explicitly noted that “[t]he trial court retains the total discretion to include or exclude this type of evidence.” 140 CONG. REC. H5437-03, H5438 (daily ed. June 29, 1994) (statement of Rep. Kyl) (quoted in United States v. Enjady, 134 F.3d 1427, 1431 (10th Cir. 1998)); see also United States v. Sumner, 119 F.3d 658, 662 (8th Cir. 1997) (quoting a statement by Rep. Molinari: “In other respects, the general standards of the rules of evidence will continue to apply, including . . . the court’s authority under Evidence Rule 403 to exclude evidence whose probative value is substantially outweighed by its prejudicial effect.” Floor Statement of the Principal House Sponsor, Representative Susan Molinari, Concerning the Prior Crimes Rules for Sexual Assault and Child Molestation Cases, 103 CONG. REC. H8991-92 (daily ed. Aug. 21, 1994)). Others were not so sure. See JUDICIAL CONFERENCE OF THE U.S., REPORTS
Rule 403 appears in other sections of Rules 413-414, as occurs with Rule 609. Nevertheless, courts unanimously read Rule 403 into the text of Rules where it seemingly does not apply, and do so for reasons of basic fairness. As the Tenth Circuit explained in *United States v. Enjady*, the adoption of Rules 413-414 “without any exclusion of or amendment to Rule 403 makes Rule 403 applicable, as it is to others of the rules of evidence.” There is no exclusion of Rule 403 in the text of Rule 609(a)(2), and the principle applies equally to it.

The courts that wrestled with Rules 413-414 made clear that their application of Rule 403 was rooted in due process considerations. Criminal defendants opposing the new rules argued persuasively that the highly prejudicial evidence of prior sex crimes violated fundamental fairness. The courts acknowledged a serious problem, but noted that the Supreme Court had left open whether violations of historical character evidence protections constituted a due process violation, and were not necessarily convinced that the unfairness fell within the limited category of infractions that violate fundamental fairness. However, even the courts that acknowledged a due process problem were satisfied that the discretionary power of the judge under Rule 403 addressed any due process concerns. These cases clarified that Rule 403 served as a guarantor of due process.

The potential prejudice in 609(a)(2) cases where the prior conviction is similar to the crime charged is equally extreme and undermines the basic fairness of the trial. The remedy, application of Rule 403, should be the same.

E. **Respect for the District Court’s Function Mandates that Judges Retain the Authority to Engage in Some Balancing, Even Regarding Convictions Involving Dishonesty or Falsehood**

Oftentimes, a conviction for falsity or dishonesty will seem probative, maybe even highly probative, of the accused’s
credibility; but that does not mean the consideration is over. The trial court must have the discretion to examine the other side of the Rule 403 balance—the nature and extent of unfair prejudice. Traditionally, the district court judge serves a vital role as evidence screener. The trial court is in the best position to make the context and fact-driven determination of how probative versus how prejudicial the prior conviction will be. As argued above, requiring the admission of prior convictions and disallowing any judicial oversight is unfair to the accused. In addition, it strips the trial court of a major judicial function.

This concern with the judge’s role was the basis of Judge Tuttle’s dissent in United States v. Toney,66 which explained that, “the purpose of rule 403 was to provide judges some flexibility in cases where the possibility of prejudice is extremely great.”67 In analyzing the facts of the case Judge Tuttle observed:

It would be hard to imagine evidence more prejudicial, in a trial for mail fraud, then [sic] the defendant’s prior conviction for mail fraud. I would suggest that the probative value of a conviction involving dishonesty is substantially outweighed by the danger of unfair prejudice to the defendant, when the prior conviction concerns the same kind of offense as that for which the defendant is being tried.68

He therefore concluded that “[a] judge should not be prohibited from excluding this evidence by a rigid holding that rule 403 can never be applied to rule 609(a)(2).”69

Allowing a Rule 403 balance reflects the natural function of the trial judge who will probably be holding various hearings to rule on evidence anyway. It requires no special extra administrative energy, but reinforces the quintessential role of the judge as screener and gatekeeper. As the Supreme Court explained in United States v. Abel,70 “[a] district court is accorded a wide discretion in determining the admissibility of evidence under the Federal Rules. Assessing the probative value of [the proffered evidence], and weighing any factors counseling against admissibility is a matter first for the district court’s sound judgment under Rules 401 and 403. . . .”71 That discretion is central to and inherent in the judicial role, and it

---

66 615 F.2d 277 (5th Cir. 1980) (Tuttle, J., dissenting).
67 Id. at 283.
68 Id. at 283-84.
69 Id.
71 Id. at 54.
appears in many aspects of the judge's labors. The judicial duty to secure a fair trial devoid of substantial and unnecessary unfair prejudice can be fulfilled with a creative reading of Rule 609(a)(2).

V. JUDICIAL CONTROL OF THE EXTENT OF IMPEACHMENT UNDER RULE 609(a)(2)

For reasons of fairness, respect for the structure of the rules, and the power of the judge, Rule 403 balancing must apply to Rule 609(a)(2). If this argument buckles under the weight of the authority opposed to it, however, the very least trial judges can do is limit the way the impeachment is proved. If, indeed, a judge must allow the prosecutor to impeach an accused charged with a prior conviction for wire fraud, then perhaps the damage can be limited by prohibiting the admission of the specific nature of the prior crime. The jury could be informed that the accused had been convicted of a crime involving dishonesty, but not exactly what that crime was. Some states have done just that, allowing the fact of a prior conviction, but not the details of the prior crime.

72 There is a long tradition of the judge as gatekeeper. Judges perform that role in screening expert testimony. See Gen. Elec. Co. v. Joiner, 522 U.S. 136, 148 (1997) (Breyer, J., concurring) ("[N]either the difficulty of the task nor any comparative lack of expertise can excuse the judge from exercising the 'gatekeeper' duties that the Federal Rules . . . impose—determining, for example, whether particular expert testimony is reliable and 'will assist the trier of fact,' Fed. Rule Evid. 702, or whether the 'probative value' of testimony is substantially outweighed by risks of prejudice, confusion or waste of time, Fed. Rule Evid. 403."). Even where the rules of evidence do not apply, the judge retains a gatekeeper role. Judges act as gatekeepers under the Federal Death Penalty Act. See United States v. Pepin, 514 F.3d 193, 204 (2d Cir. 2008) ("[U]nder the FDPA [s]tandard, judges continue their role as evidentiary gatekeepers and, pursuant to the balancing test set forth in § 3593(c), retain the discretion to exclude any type of unreliable or prejudicial evidence that might render a trial fundamentally unfair."); United States v. Roman, 371 F. Supp. 2d 36, 43 (D.P.R. 2005) ("Although capital sentencing proceedings are released from the strictures of the Federal Rules of Evidence, see 18 U.S.C. § 3593(c), the trial judge retains his traditional role as gatekeeper of constitutionally permissible information, and must accordingly exclude any unreliable or prejudicial information that might render a trial fundamentally unfair.").

73 See State v. Geyer, 194 Conn. 1, 16 (1984) ("The defendant's character, from which the jury might draw an inference of dishonesty, would thus be sufficiently impugned without the extraordinary prejudice that sometimes follows when the prior crime is specifically named."); State v. Shepherd, 94 Idaho 227, 230 (1971) (requiring balancing by the judge to admit the specific nature of the prior conviction); cf. State v. Brunson, 132 N.J. 377, 391 (1993) (limiting evidence of conviction to "the degree of the crime and the date of the offense but excluding any evidence of the specific crime of which defendant was convicted"); State v. White, 43 Wash. App. 580, 586 (Ct. App. 1986) (citing State v. Jones, 101 Wash. 2d 113, 121 (1984)) (placing the decision whether to allow specifics about the prior crime in the hands of the district court). But
From where would the judge, bound by the current text of 609(a)(2), derive authority to follow these state courts' lead? If the answer is Rule 403, we would seemingly be right back where we started, and this proposal could not serve as a next-best method to avoid the intolerable unfairness of impeaching the accused with evidence of a similar prior crime. However, it cannot be true that no part of Rule 403 applies to admission of impeachment for a prior crime of dishonesty. Certainly, the prosecutor cannot, in the course of proving a prior conviction involving dishonesty or false statement, make outrageous comments about “bloodsucking hucksters who prey on the elderly” or call weepy victims to prove the prior fraud. Such tactics would be prohibited even if the specific nature of the crime is admissible. Hence, some residual power vested in the court under Rule 403 and Rule 611, which controls the order and mode of proof, support the inherent ability of the trial court judge to manage the evidence. Limiting the prior conviction for an act of dishonesty to the fact of such conviction, without naming or describing the exact crime, would mitigate the immense prejudice against the accused. In addition, this

see People v. Van Dorsten, 298 N.W.2d 421, 421 (Mich. 1980) (“It is improper to impeach defendant by telling jury only of existence of unnamed prior felony convictions, without providing names of the offenses, since it is the nature, rather than the fact, of prior felony conviction which jury is to use in its evaluation of credibility.”).

State v. White, 43 Wash. App. 580, 586 (Ct. App. 1986) (citing State v. Jones, 101 Wash. 2d 113, 121 (1984)) (“[T]he determination of whether to name or not name the prior convictions introduced for the purposes of impeachment should rest with the discretion of the trial judge as an additional aspect of the ultimate determination that the prejudicial effect of the evidence on the defendant does not outweigh its probative value.”).

Rule 611, entitled “Mode and Order of Interrogation and Presentation” in section (a) provides: “The court shall exercise reasonable control over the mode and order of interrogating witnesses and presenting evidence so as to (1) make the interrogation and presentation effective for the ascertainment of the truth, (2) avoid needless consumption of time, and (3) protect witnesses from harassment or undue embarrassment.” The advisory committee note explains:

[Rule 611] restates in broad terms the power and obligation of the judge as developed under common law principles. It covers such concerns as whether testimony shall be in the form of a free narrative or responses to specific questions, the order of calling witnesses and presenting evidence, the use of demonstrative evidence, and the many other questions arising during the course of a trial which can be solved only by the judge’s common sense and fairness in view of the particular circumstances.

Fed. R. Evid. 611 advisory committee’s note (citations omitted).

cf. 4 Weinstein’s Federal Evidence, supra note 43, § 611.03(4)(a), at 611-42.2 (explaining that under Rule 611, a district court balances “the factors of prejudice, confusion, and delay against the probative value of the testimony” in deciding whether to limit cross examination).
proposed limitation fits the pattern of *Old Chief v. United States*, where the Court held that in a prosecution for being a felon in possession of a gun, a prosecutor was not entitled to admit details of a prior felony if the accused would stipulate to his status as a felon. Although *Old Chief* arose in the context of proving a status, it is similar in posture to the Rule 609(a)(2) scenario where the proof of a fact is collateral to the underlying accusation and presents the danger of inviting the jury to make an impermissible propensity inference. *Old Chief* can operate by analogy only because, of course, it was a Rule 403 case. Nevertheless, its focus on the process of proving an admissible fact for a collateral purpose is instructive.

**CONCLUSION**

My sensible proposal is that all evidence must be screened for unfair prejudice, and where the unfair prejudice substantially outweighs the probative value of such evidence, the trial court must, out of fairness for the accused and respect for the trial court, possess the discretion to disallow it. This principle applies even to impeachment under Rule 609(a)(2) for felonies of dishonesty and false statement. I do not propose any special, more favorable tests for the accused (as appears in Rule 609(a)(1)), but argue that a basic Rule 403 balance, which tends to admit all but the most unfairly prejudicial evidence, is absolutely fundamental to due process. Indeed, in other contexts, it has been heralded as the guarantor of due process. Although courts will doubtlessly conclude that most convictions regarding dishonesty and false statement are highly probative and do pass the balancing test of Rule 403, there are clearly some cases where such a test cannot be passed, especially where the accused is the witness. In such cases, the fairness of the trial is put in question without a Rule 403 balance.

Given that there is currently no balancing, and hence no point in making a motion in limine to exclude evidence of convictions for dishonesty or false statement, it is impossible to know how many of the accused forgo their right to take the stand in their own defense. This discouragement from taking the stand is seriously troubling because the accused could

---

76 529 U.S. 172, 201 (1997).
77 See United States v. Toney, 615 F.2d 277, 283 (5th Cir. 1980) (Tuttle, J., dissenting) (“[C]onvictions involving dishonesty could be excluded only upon a strong showing of overwhelming prejudice to the defendant.”).
present useful information and juries tend to be more conviction-prone if the accused does not testify.

Finally, why did I choose this sensible idea to honor Margaret Berger? Professor Berger is justly famous for her work on scientific evidence and expert testimony, but has not written extensively on character evidence. However, in her immensely important treatise and her various *Science for Judges* articles Professor Berger, undoubtedly influenced by the model of her teacher and collaborator, Judge Jack Weinstein, evinces faith in the trial judge and the importance of the judicial “gatekeeper” role. The tenor of Professor Berger's remarkable career reflects a genuine trust in the capacity and good sense of trial judges, and she has invested her time and talent into educating them. Where judges are wrong-headed, Professor Berger has never shied away from challenging mistaken rulings. In her writings on evidentiary error, the confrontation clause, and expert testimony, Professor Berger also has championed the rights of the accused. Even though I am challenging a statement in her own treatise, what better way to honor her than to propose a reinterpretation of a rule that restores discretion to judges to preserve fairness for criminal defendants?

78 2 WEINSTEIN'S FEDERAL EVIDENCE, *supra* note 43, § 403.02(01)[a], at 403-6 (“The one instance in which Rule 403 does not apply is in ruling on the admissibility of convictions pursuant to Rule 609(a)(2).“).
Reflections on Teaching Evidence with an Audience Response System

Roger C. Park†

I am flattered to have been invited to contribute to this festschrift in honor of Margaret Berger. I have great respect for her contributions to scholarship in evidence and civil procedure. Her work has often helped me with my own. She has written about law and science at the highest level, while finding time to serve as a consultant and reporter on highly significant projects that affect the law in action. She is also the co-author of a leading treatise on evidence and a leading casebook. Among her works is an article on forensic evidence, which I teach in my basic Evidence course using the response system that I will describe in this article.

I started using an audience response system in my spring 2009 Evidence course. I project a question on a screen, and students signal their answers using “clickers.” When enough students have answered a question, their answers are automatically displayed with an accompanying graph. The screen display shows aggregate answers without revealing which student gave which answer. For an example, see Figure 1.

† James Edgar Hervey Professor of Law, University of California, Hastings College of the Law. I would like to thank Eric Noble, Eugene Wu, and Jeremy Hessler for helping me to set up and use the audience response system, and the participants in the Spring 2009 UC Hastings Teaching Discussion Group for their helpful comments.

2 Among her other activities, Professor Berger has served as Reporter to the Advisory Committee on the Federal Rules of Evidence, as a consultant to the Carnegie Commission on Science, Technology and Government, and as a reporter for a working group of the National Commission on the Future of DNA Evidence.


4 The “clickers” were handheld radio transmitters that sent a signal to a radio receiver that was plugged into the USB drive of my computer. The “clicker” hardware may become obsolete as web-based systems that receive signals from students laptops and cell phones become available. See infra notes 20-21 and accompanying text.
After projecting the answers to the initial question, I follow up in class by asking other questions, either using screen projection or asking the follow-up questions orally, and calling on individual students to respond. Sometimes I ask students to discuss answers with their neighbors and re-answer a question.

The PowerPoint platform was the foundation for the application that I used to poll audience responses. PowerPoint teaching was a new experience to me. In earlier years, I avoided it because I thought it would tie me down too much. Also, Edward Tufte’s essay on the Cognitive Style of

---

4 After the question presented in Figure 1, I projected the following question and asked students to signal their answers:

The statement is not hearsay if it is offered to prove—

1. She sent the marijuana
2. The package contained marijuana
3. Buzzy believed the package contained marijuana
4. More than one of the above

When the answers to my initial questions are debatable, I usually call on a student to defend one of the answers instead of projecting follow-up questions.

5 The add-on application was TurningPoint. See description, infra note 14 and accompanying text.
PowerPoint made me worry that there might be something inherently corrupting about the way the system encouraged presenters to organize their thoughts, and that a system developed for pitching “power points” in the business world might not be right for academia. The addition of an audience response system to PowerPoint turned the tide for me. I have always believed in active learning; it worked for me as a law student, and early-career research on teaching and learning reinforced my belief in it. Another reason I tried the system was that laptops had become ubiquitous, causing some students to be virtually not present during class. I was reluctant to ban laptops, but I welcomed anything that would compete with them. Finally, the system provided extra feedback to students. I was not disappointed. So far as I could tell, students paid close attention to the questions and worked actively on solving them. The use of laptops for passive note-taking (or worse) seemed to decrease. During the answer pauses, students looked at the projector screen instead of their laptop screens. When I asked them to talk to each other about the question, they did so vigorously.

---


7 In Tufte’s words,

The standard PowerPoint presentation elevates format over content, betraying an attitude of commercialism that turns everything into a sales pitch. . . . PowerPoint’s pushing style seeks to set up a speaker’s dominance over the audience. The speaker, after all, is making power points with bullets to followers. Could any metaphor be worse? Voicemail menu systems? Billboards? Television? Stalin?

Tufte, PowerPoint is Evil, supra note 6.


9 For further discussion, see Caron & Gely, supra note 8, at 554-58.

10 Id. at 564-65.
The students also liked the audience response system. In an anonymous poll administered in my class, 97% of the respondents agreed with the proposition that “[c]lickers have been beneficial to my learning.”

![Figure 2](image)

**FIGURE 2**

The audience response system changed the timing of class interaction. When I asked a question, I paused for students to respond. As students answered the question, the screen display revealed how many had answered, though it did not reveal the breakdown of answers into categories until getting a further signal from me. I usually waited until 50 or 60 students answered the question. (Waiting for 60 answers meant about a 90% response from students armed with

---

11 See Figure 2. The answers were anonymous, but of course they might have been influenced by the fact that I administered the survey. Sixty-eight students were in the class the day I did the survey and I got 58 responses. Four of them couldn’t respond because they didn’t bring their “clickers,” and six of them had “clickers” but didn’t answer in time. After submitting this essay, I used the “clicker” system again the following semester and asked the same question to my evidence class on November 4, 2009. Seventy-six students attended that class. Seventy-two brought their “clickers.” Sixty-eight responded to the assertion “[c]lickers have been beneficial to my learning.” The answers were as follows: 66% strongly agreed, 31% agreed, 3% were neutral and 0% disagreed or strongly disagreed.
“clickers.”) My teaching assistant timed the length of the answer pauses in five of my early classes. The results are in Figure 3. In those classes, I asked 56 questions using the system, or just over 11 questions per teaching hour. The average pause-for-response time was about one minute per question. The curve was skewed right; sometimes I gave the students two or three minutes to answer. These long pauses occurred when I asked students to discuss answers with their neighbors.

**Answer pauses**

(This chart shows answer response times in five 50-minute classes in which a total of 36 ARS questions were asked)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>55 seconds</td>
</tr>
<tr>
<td>Mean</td>
<td>65 seconds</td>
</tr>
<tr>
<td>Range</td>
<td>15-165 seconds</td>
</tr>
</tbody>
</table>

**FIGURE 3**

I relished the extra thinking time the answer pauses gave me. During the pauses, I had time to arrange my notes or look at the seating chart. I could think about whether to use volunteers to provide explanations for the answer, and decide who should be called upon for follow-up questions. I could look at the students to try to figure out what they were doing. I could walk up to them to eavesdrop on their discussions, or ask whether they were prepared to answer oral questions after the pause was over.

12 Occasionally I polled the students about whether they had brought their “clickers” to class. I also asked my teaching assistant to count the number of students in the class. Typically there were four or five more bodies than “clickers.”
Admittedly, the answer pauses have the potential disadvantage of lengthening the class and cutting down on coverage. However, some of the time loss is recouped because, when students are asked to explain their answers, they are less likely to ask that the question be repeated or to request a pass because they don’t have anything to say. Also, some steps in an ordinary question-and-answer class can be skipped or minimized. When I prepare notes for a Socratic class, I often include “helpers” or follow-up questions that are designed to bring out relevant points when the student is baffled by the question. These questions are sometimes useful, but they take time. At any rate, it is a refreshing change to have students give more thought to the question before answering it. (A similar result can be obtained by assigning problems and questions ahead of time and pre-appointing the students who will give answers, but I do not often use that method. It promotes passivity among the other students, and sometimes only the “expert” students and the teacher understand the dialogue.)

In the end, I concluded that whatever drawbacks the answer pauses had, they were more than offset by the likelihood that the audience response method promoted widespread active involvement of students. The opportunity for me to have a few minutes to think and plan at midpoints in the class was a nice side effect.

The audience response system also facilitated a “talk to your neighbor” approach to teaching. Of course, it is always possible, under any system, to ask students to solve problems together, either before or during class. But the audience response system adds something because it tells the teacher where students are having problems with a concept. When the display reveals an unexpected pattern of answers—for example, when many students choose an answer that the teacher believes is clearly erroneous—the teacher can ask students to talk the question over with their neighbors and then re-poll them.

I tried this technique because a colleague had sent me an article from *Science* reporting on the use of audience response teaching in a genetics course. The authors used an

---

13 M. K. Smith et al., *Why Peer Discussion Improves Student Performance on In-Class Concept Questions*, 323 *Science* 122 (2009). The abstract of the article reads as follows:
audience response system and asked students to talk answers over with their peers. The authors reported that this technique resulted in superior performance, compared with the standard teaching method, not only on the talked-over questions, but also on other similar questions. I did not collect data of this nature, but I did try the method and noticed that students generally made good progress in reaching the correct answer, a result that could not always be attributed to copying the majority answer (sometimes the original majority answer was wrong).

I have mentioned the “correct” answer, a concept that will not sit well with some law teachers. I might as well confess that I believe that there are analytically challenging questions about law that have one clearly superior answer. I am not afraid that by asking such questions I will be conveying the message that all legal questions have a definite answer. That battle was won long ago, and today’s students are appropriately skeptical about the certainty of the law.

In ordinary question-and-answer classes, questions that have definite right answers sometimes fall flat. It is harder to salvage student answers to one-answer questions than to questions that have many correct or arguable answers. In theory, the Socratic master should, by asking the right questions, be able to lead students to discover the right answer themselves, without just telling them what it is. But things do not always work that way. Sometimes the teacher ends up saying the equivalent of, “Well, your answer was ‘yes,’ now what’s the other answer?” The falling-flat problem can lead the teacher to avoid asking questions that have clear right and wrong answers, at least when the questions are difficult ones. If so, then common misconceptions are not examined with an active learning method; difficult concepts are explained by lecture and absorbed by passive note-takers.

When students answer an in-class conceptual question individually using clickers, discuss it with their neighbors, and then revote on the same question, the percentage of correct answers typically increases. This outcome could result from gains in understanding during discussion, or simply from peer influence of knowledgeable students on their neighbors. To distinguish between these alternatives in an undergraduate genetics course, we followed the above exercise with a second, similar (isomorphic) question on the same concept that students answered individually. Our results indicate that peer discussion enhances understanding, even when none of the students in a discussion group originally knows the correct answer.
The audience response system facilitates asking difficult analytical questions that have a clearly preferable answer. If the teacher sees that many students have gone astray, she can ask students to discuss the question with a neighbor and then re-poll. Having done her duty to facilitate active learning, after re-polling she can choose just to explain the answer, rather than trying to derive it from further questions.

That does not mean that the teacher should stick only to questions that have clear, correct answers. The audience response system is not a graded test (or at least it need not be one). The teacher is free to ask questions that do not have a clear, correct answer. These conversation starters can facilitate later class discussion or lead to sub-questions, exploring justifications for competing answers.

Of course, the fact that the question-asker has to offer the students a set of predetermined answers can be a disadvantage. Some questions do not fit that approach. Standards like “What’s the holding of the case?” and “How are those two cases distinguishable?” work less well with predetermined answers. But the system works well with hypotheticals that ask the student to apply a case or statute to a new situation. Also, questions about values and objectives can easily be asked; they will lead to a free-form discussion that the teacher can continue off-screen. The fact that the teacher uses an audience response system does not mean that it must be used for every question. Anytime he wishes, the teacher can darken the screen and continue with another method of instruction.

Because there is always an answer pause of at least a few seconds, the teacher might want to avoid asking extremely simple questions to which he expects a quick answer. If the teacher wants to use a short drill of simple questions, he should consider rolling all those questions into one by listing them in a format that makes it possible for the final option to be “all of the above” or “none of the above.”

Sometimes a question with a list of alternative answers can be used as a substitute for a lecture. Suppose that the teacher wants to make sure that the students are familiar with the standard justifications for a rule, because she later plans to compare it to another rule and ask if the two are consistent. The quickest approach would be to simply list the justifications that have been advanced for the rule in a lecture. But the answer may stick better with the students if they are presented with a list of objectives and asked which ones support the rule.
That way the students evaluate each objective and are more likely to remember them.

I hope that my questions will improve from year to year, because I certainly learned more about my own questions and hypotheticals by using the audience response system than I would normally learn using a verbal question-and-answer method of teaching. The reason is that the audience response system can reveal that an unexpected answer is widely shared. I expect that many readers have had this experience when reviewing data about student answers to exam multiple-choice questions. With the audience response system, the teacher has the real-time option of probing to see why students chose an unexpected answer. This can reveal that the question was ambiguous, or that there is an unforeseen good argument in favor of one of the response alternatives. Sometimes it reveals a misconception whose roots need to be explored.

The audience response system also made me more aware of ways in which answers might turn on facts not stated in my hypotheticals. Perhaps this was an outgrowth of the fact that my questions sometimes expressly offered an “it depends” answer, or the fact that I would ask students to explain when I got unexpected responses from a significant number of students. As the year went on, I found myself trying to define what sorts of factual variations were fair game in answering the questions. For example, I deemed an “it depends” answer to be fair only when the answer depended on facts that were reasonably probable under the circumstances. For example, a student should not answer a hearsay question “it depends” on the grounds that the declarant was aware of the imminence of death, unless the other facts in the hypothetical suggested that it was reasonably likely.

I have only used one audience response system, TurningPoint,\textsuperscript{14} so I cannot compare it to other systems. That being said, it seemed to me that TurningPoint was relatively “unbuggy” and easy to use. I enjoyed making my slides.

TurningPoint does have a few regrettable features, but these can easily be cured by changing default settings or by making template slides. The attached footnote describes two of

\textsuperscript{14} For the manufacturer’s description, see Student Response Solutions, Turning Technologies, http://www.turningtechnologies.com/studentresponsesystems/studentresponsesolutions/ (last visited Feb. 27, 2010). I have no relationship with the makers of TurningPoint or PowerPoint other than as a customer for their software.
the features I changed. I cannot resist commenting in text about one feature that particularly annoyed me. The feature exists for all types of questions and all ways of displaying the answers, but I will illustrate with an example involving a question that has the answer choices “yes” and “no,” and whose answers are displayed using a bar graph. If 67% of the students give the “Yes” answer and 33% give the “No” answer, then the bar representing the “Yes” answer will be twice as high as the one representing the “No” answer. That’s fine. But suppose the vote is closer, so that the number of students giving the “Yes” answer is only slightly larger than the number giving the “No” answer—for example, 52 to 48. In the bar graph display, the bar representing the “Yes” answer will still be about twice as high as the one representing the “No” answer! This led to laughter when it occurred in my class, because law students, whatever their quantitative shortcomings, know that 52 is not twice as much as 48. I assumed that this feature was a flaw in the application, and was surprised to find that it was deliberate. TurningPoint’s explanation was that if the difference was not exaggerated, the audience might not realize that one answer got more votes than the other. Fortunately, this ghastly feature can be turned off, and when it is off, the answer graphs will display actual proportions. In other words, the default display intentionally makes small differences look the same as big differences. If this way of displaying data pervades the business world, it may explain our current financial crisis.

The TurningPoint system is capable of gathering individual and group data about student answers. Each “clicker” has a different radio signature, and by keeping track

15 (1) The space allotted for questions is center-aligned, so that the questions are formatted like a title page. Left alignment (such as that used in the text of this essay) can be achieved by changing the default settings. (2) The slides allot too little space for questions and too much space for answers. This can be changed by making a template slide and copying it instead of making new question slides by clicking on “new slide” on the toolbar.

16 To be precise, the explanation given for showing charts that are not proportionate to answer choices is that the method “will clearly identify the answer choice that received the majority of responses.” (The quote comes from a pop-up window that appears if a user hovers the cursor over the “Auto Scale Charts” option under “Chart Settings” in TurningPoint version 4.0.0.8224.)

17 To do so, (1) Choose “Tools” (in the TurningPoint menu), (2) Choose “Settings,” (3) Make sure “Presentation” is highlighted, (4) Change the setting for “Auto Scale Charts” to FALSE, and Choose “Done.” I am indebted to Hastings’ IT Director, Eric Noble, for this guidance.
of who had which “clicker” a teacher could collect data about individual student answers. Demographic data could also be collected; for example, you can compare answers of male and female students, or compare the answers of students who did well on other questions with those who did poorly. The system can also be an attendance checker.

I decided not to use these features. I knew that data collection would cause student concerns, and I was not sure what I would do with the information anyway. So I set the data collection feature on “anonymous.” (Teachers who want to take an extra step to reassure students can also ask them to swap “clickers” with their neighbors when answering sensitive questions, or by not collecting information about which “clicker” went to which student in the first place.) To give students an extra feeling of safety, I decided not to give credit for class participation, and I announced that decision on the first day of class. I am happy to report that the participation of typical class members has been better than in years in which I did give grade credit for participation.

I posted my slides on my class website after using them. I hope that this encouraged students to focus on thinking about the questions instead of trying to write them down verbatim, since they could always cut and paste the questions from the website later. To help the students find particular slides for cut-and-paste, I numbered the pages on the slides I displayed in class.

I found it helpful to have a teaching assistant, though I am sure many teachers could do without one. I hired a member of the class to set up the computer and projection equipment each day, summon help in emergencies, format the slides for posting on the web, and collect data such as that displayed in Figure 2. I made the slides myself in my office, often tinkering with them until just before class as I rehearsed. With the projector in place and working as a regular element of the class, it of course was easier to do other things that required it, such as showing video clips.

Using TurningPoint meant that I was also using PowerPoint. I learned to like PowerPoint more than I had expected, though I never figured out how best to handle large chunks of text. One option is to explain the display, but not to

---

The Director of IT at Hastings handed out the “clickers” in my class, noting which student had which “clicker.” Students signed an agreement promising to pay the school $35 if they failed to return the “clicker.”
read it, which divides the students’ attention because they are simultaneously trying to read the display.19 Another is to read the slide, which is deadly and which research suggests is also distracting, because students will monitor the reading.20 The third is to step aside and be a stagehand, which does not seem better than simply telling the students to turn to a page in the book and read it to themselves. So I gradually learned to avoid large text displays. Thus, when lecturing, I would display short “bullet points” that raised topics, then I would explain the topics orally. I spent more time in question-and-answer mode than I did lecturing.

When I was conducting a dialogue about the case, I would sometimes just display the case name, court, date and page number, hoping that information would be a useful reference for confused or drifting students. At other times I would add a line or two about the facts of the case. This text might describe the gist of a statement whose hearsay status was in issue, or names and postures of parties and witnesses. For one case, I used a simple drawing of an auto mishap to summarize its facts. I also used PowerPoint to display diagrams, graphics, charts and photos.

On the question slides, I would sometimes try to keep the text display short by stating part of the hypothetical orally, only putting a reminder of its gist on the slide. More often, I made an exception to my rule that text displays should be short, and displayed the entire hypothetical so that students could refer precisely to what was being asked in formulating their answers. When I did set out the facts of a hypothetical at length on the slide, I would either remain silent while students read it, or read it aloud myself word for word. I generally avoided other commentary that could have distracted students while they were trying to read the facts of the hypothetical. When I was using a hypothetical printed in the casebook, I did not display its facts on the slide. The slide displayed only the page number of the hypothetical and its bottom-line question (e.g., how should the judge rule?). I did things this way because when I intend to use a hypothetical printed in the casebook as a question in class, I always assign the question in the syllabus and implore the students to answer it before class; I wanted the students to look at these hypotheticals in their casebooks.

20 Id. at 47-48.
instead of on slides, because I was hoping that they had annotated their casebooks with answers, or at least that the question would seem more familiar to them if they looked at it again in the book. On pure-text slides that contained no questions, sometimes I thought it made sense to put up a fairly long excerpt from a statute, for the purpose of using PowerPoint features to highlight and discuss specific words or phrases in the statute. But generally, as the year went on, my pure-text slides got shorter and shorter.

For the TurningPoint system that I used, students signaled their answers with handheld “clickers” that sent a radio signal to a receiver plugged into the USB drive of my computer. Alternatives to handheld “clickers” are becoming available in the form of web-based systems that take answers from students’ laptops or cellphones. Perhaps some web-based systems will be free, or included as an element in the school’s subscription to some other service, such as Westlaw or CALI. At the time of this writing, Westlaw’s TWEN system offers a “customized polling” option that has some of the features of the audience response system that I used, though the current TWEN system seems more suited for pre-class polling than for live in-class use. Although the use of handheld “clicker” hardware is not essential to an audience response system, having dedicated hardware for signaling may have advantages. There is no delay for student log-ins, and “clickers” can be used by professors who have banned laptops.

I was pleased with my experience with an audience response system. It provided active learning and feedback without much time or trouble. Judging from the results of the attitude surveys that I have described, students also liked this method of teaching.

---

21 TurningPoint recently announced that it would be offering a web-based system. See Andrea Lawn, Press Release, Turning Technologies Unveils ResponseWare™ Web—a Web-based Polling System, SCYONLINE (March 31, 2009), available at http://blog.svconline.com/briefingroom/2009/03/31/turning-technologies-unveils-responseware%E2%84%A2-web%E2%80%94a-web-based-polling-system/.

22 Faculty members at Westlaw-subscriber schools can try out the polling feature by going to http://lawschool.westlaw.com/twen (last visited April 23, 2009). Click on “create new course” and follow the instructions for creating a course. After setting up the course, click on “customized polling” and “create a poll.”
Thoughts About *Giles* and Forfeiture in Domestic Violence Cases

*Myrna S. Raeder*

**INTRODUCTION**

In choosing a topic for this festschrift celebrating Professor Berger’s venerable career as an evidence professor and scholar, I decided that since there were likely to be many contributions in the field of scientific evidence, I would write instead about a recent Supreme Court pronouncement concerning the Confrontation Clause. Nearly twenty years ago, Professor Berger explicated a prosecutorial restraint model of Confrontation Clause analysis. She has also authored Supreme Court *amici* briefs in *Idaho v. Wright* and *Lilly v. Virginia.* While I am sure she will not agree with all of my views about *Giles v. California,* I know that she will enjoy reading about the topic.

*Crawford v. Washington* was a watershed case in Confrontation Clause jurisprudence, which rejected the reliability test articulated in *Ohio v. Roberts* in favor of a
“testimonial” approach. Justice Scalia’s originalist reading of the Confrontation Clause now requires exclusion of testimonial hearsay if the declarant does not testify at trial, unless the prosecution can demonstrate the declarant is unavailable and there was a previous opportunity to cross-examine her.\(^7\) In the wake of *Crawford*’s reshaping of Confrontation Clause analysis,\(^8\) lower courts were left reading tea leaves to discern how to apply this new framework, particularly in domestic violence cases where the complainant typically refuses to testify, or has been permanently silenced by her abuser.

The rub then, was to figure out what is testimonial and what is not in a domestic violence setting, a matter not inconsequential given that *Crawford* led to wholesale dismissals of domestic violence charges, and increased the number of trials of defendants who refused to plead guilty optimistically predicting that the absence of the complainant would result in their acquittal.\(^9\) The domestic violence advocacy community hoped the next major Confrontation Clause case, *Davis v. Washington*,\(^10\) would opt for a domestic violence exception to general Confrontation Clause analysis. Failing this, they urged a contextual view of domestic violence that recognized the patterned nature of such abuse, and, as Professor Tuerkheimer has suggested, its temporal aspects.\(^11\) *Davis* consolidated two separate cases in which the complainants failed to testify at trial, one involving a 911 call introduced against Davis, and the other involving statements made to officers at the defendant Hammon’s home by his wife.

While *Davis* was supposed to fill in the blanks, its bright-line test for determining if a statement is testimonial has been criticized as unworkable by Justice Thomas,\(^12\) and does not answer many of the hard questions posed when trying to define testimonial statements.\(^13\) *Davis’* circuitous definition

---

\(^7\) 541 U.S. 36 (2004).
\(^8\) Id. at 53-54.
\(^12\) Id. at 842 (Thomas, J., concurring in part and dissenting in part).
of testimonial statements, essentially saying statements are testimonial, unless they are not, predictably led to even more confusion about applying the testimonial approach to dual purpose statements—statements in which the victim sought help from law enforcement, but which also arguably serve as after-the-fact descriptions of criminal behavior.

Unhappy with the mixed result in *Davis*, which reversed Hammon’s conviction and affirmed the conviction of Davis, victims’ advocates pinned their hopes on forfeiture by wrongdoing as a way to ensure the admission at trial of testimonial statements made by absent domestic violence victims. In other words, if the defendant’s own conduct caused the absence of the declarant at trial, then the defendant had forfeited his right to object to the admission of her testimonial hearsay. *Crawford’s* oblique reference to forfeiture was reiterated in *Davis*, emboldening advocates to argue that forfeiture was required when the defendant killed a victim who had previously made statements to the police identifying him as her abuser. A number of courts readily agreed and did not require any demonstration that the defendant actually intended to cause the victim to be absent as a witness at trial to establish forfeiture. Instead, they only required a preliminary fact showing that the defendant murdered the victim thereby causing her unavailability.

*Giles v. California* was typical of such cases. The defendant was charged with murdering his ex-girlfriend, Brenda Avie. No one saw the incident, although the defendant’s niece heard the victim call “Granny” several times. Brenda’s wounds were consistent with her having been the victim, rather than the assailant, but the defendant claimed she was insanely jealous, had threatened him, and had previously pulled a knife on someone else. The defendant further claimed he got a gun in self-defense and shot her accidentally. To rebut this, the prosecution offered the victim’s statements to an officer describing a previous incident of

---

14 *Davis*, 547 U.S. 813.
15 Id. at 833.
18 Id. at 2681.
19 Id.
20 Id.
domestic violence, where she claimed the defendant was jealous and threatened to kill her if she cheated on him.\textsuperscript{21} The court admitted these statements pursuant to a hearsay exception that allowed admission of threats of bodily harm, and the prosecution argued that the defendant forfeited his right to confront the victim.\textsuperscript{22}

\textit{Giles} provided the Supreme Court with an opportunity to define the boundaries of forfeiture. Instead, like \textit{Crawford} and \textit{Davis}, \textit{Giles} left open more questions than it answered, further muddying the waters in domestic violence cases, and offering no guidance about interpreting forfeiture in child abuse litigation. \textit{Giles} explicitly held that constitutional forfeiture required an intent to deter a witness from testifying,\textsuperscript{23} in essence viewing such forfeiture as a sanction for witness tampering, not simply an equitable remedy for preventing the victim from testifying. It reached this conclusion by delving into the historic record. Again, like \textit{Crawford} and \textit{Davis}, the majority decision was written by Justice Scalia. Again, he utilized an originalist approach, and again his reading of history was challenged by other Justices. While the holding appears to be a blow to prosecutors trying to admit statements of dead victims, dicta in three of the five opinions in \textit{Giles} taken together implies that potentially all of the Justices are inclined to treat evidence of forfeiture flexibly in domestic violence murder cases. Moreover, the case offers a number of other hints about how individual judges think about the evolving testimonial framework. \textit{Giles} also raises questions about the nature and function of originalism in constitutional analysis.

I. ORIGINALISM

\textit{Giles} began with a reprise of \textit{Crawford}’s admonition that the Confrontation Clause was “most naturally read as a reference to the right of confrontation at common law, admitting only those exceptions established at the time of the founding.”\textsuperscript{24} I have previously argued that originalism is bound to silence voices that were not heard in 1791, when the Sixth Amendment was adopted. At that time, domestic violence was

\begin{flushleft}
\textsuperscript{21} Id. at 2681-82. \\
\textsuperscript{22} Id. at 2682. \\
\textsuperscript{23} Id. at 2692. \\
\textsuperscript{24} Id. (quoting Crawford v. Washington, 541 U.S. 36, 54 (2004)).
\end{flushleft}
neither understood, nor criminalized, and the power of chastisement was recognized by the rule of thumb. Moreover, surviving wives would have been disqualified as witnesses against their husbands under the spousal disqualification doctrine. As the Court in *Trammel v. United States* observed:

This spousal disqualification sprang from two canons of medieval jurisprudence: first, the rule that an accused was not permitted to testify in his own behalf because of his interest in the proceeding; second, the concept that husband and wife were one, and that since the woman had no recognized separate legal existence, the husband was that one. From those two now long-abandoned doctrines, it followed that what was inadmissible from the lips of the defendant-husband was also inadmissible from his wife.

Therefore, it is not surprising that the cases referred to in Giles’ historic analysis are murder cases. Generally, trials looked nothing like they do today: neither the defendant, nor the interested witnesses testified, there was virtually no hearsay, no police officers to investigate or testify, and the trials were typically very short. Thus, using a historic approach towards confrontation, which has been described as a trial right, is likely to produce incongruous results. While originalists commonly maintain that historic rights can be applied to unforeseen situations, it is somewhat different to untether the common law right of cross-examination from its 1791 context and then apply it to a world that not only uses substantially different evidentiary rules and investigative practices, but also reflects totally different social and political judgments about domestic violence.

Further, the focus on the historic record has not resulted in consensus about the common law jurisprudence. Not only have several justices argued with the majority’s views about the nature of the confrontation right in 1791, but academics and legal historians have also weighed in on

---


27 Id. at 44.


opposite sides of the issue. Thomas Davies, a legal historian, has argued that hearsay exceptions other than dying declarations were invented only after the framing, and that the framers would likely have condemned the expansion of hearsay exceptions. Therefore, it is odd to posit a theory of forfeiture for hearsay that would not have been admitted. I have often wondered why this reality has not produced arguments that hearsay exceptions enacted legislatively after 1791 are per se testimonial, since they are creations of governmental action that admit uncross-examined statements, just like those created by law enforcement. Moreover, as Justices Souter and Ginsburg indicated in their concurrence in 


33 Giles, 128 S. Ct. at 2695.

34 Myrna S. Raeder, Domestic Violence, Child Abuse, and Trustworthiness Exceptions After Crawford, CRIM. JUST. 24, 31 (Summer 2005).

35 Giles, 128 S. Ct. at 2691 (“it amounts to self-immolation”).

Justice Scalia claims intentionality is a historic feature of forfeiture. But the larger questions are why has originalism overtaken Sixth Amendment Confrontation Clause analysis at all, and does it have any significance for interpretation of other criminal procedural and trial rights? Arguably, if a strict originalist approach were applied to other constitutional guarantees it would likely strip away many of the rights criminal defendants now enjoy. For example, Washington v. Texas, which rejected the traditional witness disqualification rule under its interpretation of the Sixth Amendment right to compulsory process, quoted a Supreme Court decision from 1918 for the proposition that modern criminal procedure should not be governed by “the dead hand of the common-law rule of 1789.” Would the present focus on originalism discount compulsory process or other rights bound up in the right to present a defense? Similarly, Professor Dripps has suggested that the right to counsel could be affected by originalist reinterpretation. The irony of originalism in the Confrontation Clause context is that the right sounds absolute and defense-oriented, but the historic approach narrowly construes the content of the right, making it inapplicable to accusatory nontestimonial hearsay where the defendant would want to cross-examine the unavailable declarant. In essence, originalism gives with one hand and takes away with the other.

This same type of push-pull is becoming more noticeable in Fourth Amendment analysis. Professor Davies has contended that under “law-and-order originalism, the expansive police powers endorsed in contemporary search and seizure rulings are foreign to the Framers’ understanding of criminal procedure.” In addition, in the Fourth Amendment

37 Giles, 128 S. Ct. at 2683.
38 388 U.S. 14, 21-22 (1967) (internal quotation marks omitted) (quoting Rosen v. United States, 245 U.S. 467, 471 (1918) (affirming elimination of disqualification for felony conviction and reversing conviction for refusal to permit accomplice testimony)).
41 Thomas Y. Davies, The Fictional Character of Law-and-Order Originalism: A Case Study of the Distortions and Evasions of Framing-Era Arrest Doctrine in
arena, originalists appear to accept that rights such as “knock
and notice” existed at the founding, but then find that this does
not require suppression of any evidence obtained in violation of
that right, since the exclusionary rule is viewed as not con-stitutionally required, and subject to review for rea-sonableness. Even if the historic approach provides inconclusive results, originalists tend to interpret reasonableness narrowly to reject an exclusionary remedy. Thus, in the post-Crawford era, Justice Scalia has written two
majority opinions concerning the Fourth Amendment using the
historic approach that denied the application of the exclusionary rule. Ironically, the reasonableness review depends on the same type of judicial balancing that Justice Scalia denigrated when rejecting the Ohio v. Roberts reliability test in Crawford. Certainly, in the Fourth Amendment context the resort to reasonableness has resulted in shrinking protection for violations, as most recently demonstrated in Herring v. United States, where the Court asserted that it had “repeatedly rejected the argument that exclusion is a necessary consequence of a Fourth Amendment violation.”
A cynic might wonder if the resort to reasonableness as an end-run around the exclusionary rule evolved from the failure of Justices Scalia and Thomas to garner any support to jettison Miranda in the Fifth Amendment context in Dickerson v. United States.

Atwater v. Lago Vista, 37 Wake Forest L. Rev. 239, 240 (2002) (“[N]ineteenth and twentieth century[ ] American courts abandoned the Framers’ commitment to rigorously accusatorial criminal procedure as they drastically expanded police power.”).


44 See Virginia v. Moore, 128 S. Ct. 1598, 1604 (2008) (no exclusion where arrest was based on a misdemeanor driving offense that should not have resulted in arrest under state law).

45 See id. at 1608; Hudson, 547 U.S. at 600.

46 448 U.S. 56 (1980).


48 129 S. Ct. 695, 700 (2009) (exclusionary rule did not apply to police recordkeeping error).

49 Id.


Similarly, although Justice Scalia’s concurrence ostensibly favored the defense in Arizona v. Gant,” which held the search of a defendant’s car was unreasonable after he was handcuffed, arrested, and secured in a patrol car, Justice Scalia’s view of reasonableness based on historical practices would actually allow searches not previously permitted. In other words, he would also consider it reasonable to search a vehicle not only when the object of the search is evidence of the crime for which the arrest was made, but also “of another crime that the officer has probable cause to believe occurred.”

II. DEFINING THE BOUNDARIES OF “TESTIMONIALISM”

Giles provides glimpses into some of the Justices’ views about the boundaries of what I call “testimonialism.” The majority decision mentions that statements in furtherance of a conspiracy would probably never be testimonial,” repeating a similar pronouncement made in Crawford.” Professor Berger would not necessarily agree, since she suggested pre-Crawford that statements to undercover informants should be governed by her prosecutorial restraint model of confrontation to make the government’s role in shaping evidence transparent to jurors.” Given the Court’s current view of confrontation as restraining potential prosecutorial abuses, it is surprising that this issue has been summarily dismissed, albeit by dicta, in both Crawford and Giles, although this posture is consistent with a view of testimonialism that applies the confrontation right absolutely, but only to a very narrow range of statements.

Justice Scalia’s opinion did suggest that domestic violence cases were not hampered by a non-intent-based forfeiture rule because “[s]tatements to friends and neighbors about abuse and intimidation, and statements to physicians in the course of receiving treatment would be excluded, if at all, only by hearsay rules, which are free to adopt the dissent’s version of forfeiture by wrongdoing.” This reiterates Crawford’s suggestion that the Confrontation Clause does not apply outside of the law enforcement or governmental sphere,
ignoring the realistic possibility that statements will be made to private individuals as conduits for the express purpose of repeating them to the authorities. Similarly, the reference to physicians ignores any recognition that some doctors may be mandated to report domestic abuse, or that they may in some cases act as agents of the police, or in the case of forensic nurses, have roles that are decidedly dual in nature. Yet, since all of these references are dicta, written in general terms without concrete examples, courts will likely continue to divide on how broadly or narrowly they should interpret testimonial statements.

While Justice Alito voted with the majority in *Davis*, he joined Justice Thomas in *Giles*, questioning whether the victim’s statement was testimonial because of its lack of formality.\(^58\) This appears to retreat from his vote to reverse Hammon’s conviction, since lack of formality would result in virtually no statements made in field investigations or 911 calls being constitutionally protected because they were not considered the “equivalent of statements made at trial by ‘witnesses.’”\(^59\) At a minimum, under this rationale, the statements to the police in *Hammon* would be nontestimonial even if the victim’s affidavit was not, suggesting its admission would have rendered the error harmless. To date, this is still a decidedly minority view, though Justice Scalia’s cryptic reference that “we accept without deciding” that the statements were testimonial\(^60\) leaves room to speculate whether anyone else will jump ship. Ironically, the most likely candidate would be Justice Scalia based on his pre-*Crawford* view in *White*.\(^61\) However, given that he authored *Davis*, such a result appears improbable.

Yet, there may be another way of arguing these statements are not testimonial. *Davis* explained that when the Court stated in *Crawford* “‘interrogations by law enforcement officers fall squarely within [the] class’ of testimonial hearsay, we had immediately in mind (for that was the case before us) interrogations solely directed at establishing the facts of a past crime, in order to identify (or provide evidence to convict) the

---

\(^{58}\) *Id.* at 2694.

\(^{59}\) *Id.*

\(^{60}\) *Id.* at 2682.

perpetrator.\textsuperscript{62} Thus, \textit{Davis} redefined testimonial statements in the 911 or field investigation context, typical in domestic violence prosecutions:

Statements are nontestimonial when made in the course of police interrogation under circumstances objectively indicating that the primary purpose of the interrogation is to enable police assistance to meet an ongoing emergency. They are testimonial when the circumstances objectively indicate that there is no such ongoing emergency, and that the primary purpose of the interrogation is to establish or prove past events potentially relevant to later criminal prosecution.\textsuperscript{63}

However, this definition arose in a context where the statements were being used in the prosecution of the crimes arising out of the incident in question. That is not how the evidence is used when domestic violence results in murder. In other words, the previous statements are not being used to convict the defendant of the earlier assault, but rather as circumstantial evidence linking the defendant to a later crime that was not committed at the time of the earlier statement. While such statements are accusatory, and the defendant would undoubtedly want to cross-examine the declarant about them, this is not the approach to testimonialism adopted by the Supreme Court.\textsuperscript{64} Indeed, considering such statements to be testimonial in the murder case smacks of a science fiction approach to crime, suggested by Minority Report, a film in which individuals were apprehended for offenses before they committed them.\textsuperscript{65} Assuming a free will approach to criminal responsibility, accusing someone of a past crime arguably does not qualify as testimonial when used as evidence of a future crime that has not yet occurred.

This view is consistent with the Court’s explanation that the right to confrontation only arises at trial: “[t]he Confrontation Clause in no way governs police conduct, because it is the trial use of, not the investigatory collection of, ex parte testimonial statements which offends that provision.”\textsuperscript{66} The mere fact that statements arose out of a police investigation should not per se label them as testimonial when


\textsuperscript{63} Id. at 822.

\textsuperscript{64} See generally, e.g., Robert P. Mosteller, “Testimonial” and the Formalistic Definition—The Case for an “Accusatorial” Fix, 20 CRIM. JUST. 14 (Summer 2005).

\textsuperscript{65} MINORITY REPORT (Twentieth Century-Fox File Corp. et al. 2002).

\textsuperscript{66} Davis, 547 U.S. at 832 n.6 (emphasis omitted).
introduced at a trial for some other separate crime as a piece of circumstantial evidence to demonstrate motive or identity pursuant to Federal Rule of Evidence (“F.R.E.”) 404(b), or as evidence of the decedent’s state of mind, which is relevant as in Giles where the defendant claims the death was accidental or that he acted in self-defense.67

This reasoning is analogous to that in Dowling v. United States, which held that neither double jeopardy nor due process were violated by the introduction of evidence relating to a crime for which the defendant had previously been acquitted.68 Dowling was charged with a bank robbery in which the perpetrator wore a ski mask and carried a small gun.69 At his third trial for the robbery, a woman testified that Dowling entered her home along with another individual, and he wore a ski mask and carried a small gun.70 She was able to identify Dowling because she unmasked him during a struggle.71 The prosecution introduced this testimony on the theory that the mode of dress in both incidents occurring two weeks apart was similar, and the second man identified by the female witness was also involved in the robbery, thereby linking Dowling to him.72 However, Dowling had previously been acquitted of charges relating to the witness’ testimony. The Court reasoned that although the acquittal established that there was a reasonable doubt that the defendant was the masked man who entered the witness’ home two weeks after the bank robbery took place, that fact “did not determine an ultimate issue in the present case.”73 The decision recognized that the “jury’s verdict in his second trial did not entail any judgment with respect to the offenses charged in his first.”74 Moreover, under F.R.E. 404(b), similar act evidence is relevant if the jury can reasonably conclude that the act occurred and that the defendant was the actor, a standard satisfied by F.R.E. 104(b).75 In other words, it does not have to meet the preponderance

69 Id. at 344.
70 Id. at 344-45.
71 Id.
72 Id. at 345.
73 Id. at 347-48.
74 Id. at 353.
standard that typically governs the admissibility of evidence, let alone satisfy proof beyond a reasonable doubt.

Similarly, in a domestic violence murder trial the victim’s previous statement to the police is not being used to prove the offense that it describes, but simply constitutes evidence that is inferentially used to support a finding of proof beyond a reasonable doubt for a different crime, committed at a later time. Therefore, the statement should not be considered testimonial for that purpose. In a related vein, *Crawford v. Commonwealth* recently held that statements in an affidavit supporting a request for a civil protective order were not testimonial when the defendant is later charged with the declarant’s murder, since her statements were not made for the primary purpose of reporting a past event for possible criminal prosecution.76

Because *Giles* did not specifically address whether forfeiture applied to testimony initially obtained in one proceeding, but subsequently introduced at a second proceeding involving the same defendant charged with a different offense, courts are now being asked that question. For example, *United States v. Vallee*77 interpreted *Giles* “invitation . . . to the state court to explore defendant’s intent on remand” as extending forfeiture to murder trials, “provided . . . that the defendant intended to prevent the witness from testifying at an earlier proceeding.”78 Similarly, the Missouri Supreme Court interpreted forfeiture to apply to both the ongoing proceeding as well as the murder.79 This interpretation would be unnecessary if a reevaluation of such statements deemed them to be nontestimonial.

III. OPENING THE DOOR TO STATE OF MIND TESTIMONIAL STATEMENTS

Several of the statements in *Giles* and other domestic violence femicides are relevant to the decedent’s state of mind, which would be admissible under F.R.E. 803(3) when the defendant claims that the killing was in self-defense or the death was an accident or suicide.80 In such instances, the

77 304 F. App’x 916 (2d Cir. 2008).
78 *Id.* at 920-21 & n.3.
79 See *State v. McLaughlin*, 265 S.W.3d 257, 271-73 (Mo. 2008) (en banc).
80 See, e.g., *People v. Abordo Espinosa*, No. A102886, 2005 WL 941454, at *5 (Cal. Dist. Ct. App. Apr 25, 2005) (“[E]vidence of domestic violence . . . was relevant to rebut the defense’s theory that the shooting occurred in heat of passion in response to
decedent’s state of mind, typically fear of the defendant, is a window to her likely conduct in regard to the defendant. For example, the majority in *Giles* detailed the statements made to the police in the previous incident by Brenda Avie, the decedent and Giles’ ex-girlfriend, that Giles had accused her of having an affair, grabbed her, lifted her off the floor, choked her, punched her in the face, opened a folding knife, and “threatened to kill her if he found her cheating on him.”

Brenda’s wounds, which were described as consistent with defensive motions,\(^1\) would more likely be viewed as defensive if there was evidence that she feared the defendant. Similarly, the dissent characterized the defendant’s description of the victim as “jealous, vindictive, aggressive, and violent.”\(^2\) Evidence of her fear would cast doubt on the defendant’s claim that she was jealous and threatened him.

In this capacity, such statements should not be barred as testimonial, even if they were made to a police officer. First, the victim’s statements are not being used to prove the crime that the police were then investigating, so, as previously mentioned, the decedent did not expect the statements to be used in relation to the murder. Second, fear is not used to prove an element of the current crime, but only to explain the decedent’s behavior, in evaluating the defendant’s version of the facts. Third, as I have argued elsewhere, by raising a defense that makes the decedent’s state of mind relevant to the case, the defendant’s trial strategy should be viewed as waiving any confrontation challenge concerning otherwise admissible evidence that rebuts the defendant’s testimony about the decedent’s state of mind.\(^3\) On occasion, such state of mind may be nonhearsay, which clearly escapes *Crawford’s* mandate. However, even if admitted via a hearsay exception, the statements would not typically be expected by the declarant to be used at a trial involving a different incident. Of course, F.R.E. 803(3) cannot be used to prove the underlying conduct producing the declarant’s state of mind,\(^4\) so such statements are viewed as prejudicial even if instructions are given as to

---


\(^2\) *Id.* at 2681.

\(^3\) *Id.* at 2695 (Breyer, J., joined by Stevens & Kennedy, JJ., dissenting).

\(^4\) See Raeder, *Remember the Ladies*, supra note 25, at 358-60.

\(^5\) See FED. R. EVID. 803(3).
their limited use. However, if the defendant offers to stipulate to the decedent’s state of mind, this factor can be analyzed by the judge in determining if admission is unduly prejudicial under the F.R.E. 403 balancing test.

Moreover, it is arguable that defenses based on accident, suicide, and self-defense should make otherwise admissible evidence of the defendant’s state of mind admissible as well. For example, Giles claimed Brenda was jealous, and described her alleged threats and jealous statements. In contrast, she described his threats and jealousy. In this context, not only should his claim of accident and self-defense open him to admission of her statements even if testimonial, but it should also open him to the statements she made concerning his state of mind, again assuming a basis for admission other than F.R.E. 803(3). No reliance needs to be given to forfeiture hearsay exceptions to cover her statement to the police officer, since most of them would be excited utterances or fit ad hoc trustworthiness exceptions. Thus, the testimonial ban is the only barrier to their admission. If the defendant opens the door with such a defense posture, I view it no differently than any other hard strategy decision that opens the door to otherwise inadmissible evidence. For example, *Michigan v. Lucas* held that excluding evidence of defendant’s past sexual conduct with the victim for failure to comply with a rape shield’s notice provision is not a per se violation of the Sixth Amendment. If trial strategy could result in the loss of a constitutional right pre-*Crawford*, there is no reason to require a different result post-*Crawford*.

**IV. INFERRED INTENT**

The justices disagreed about the role of intent in the forfeiture analysis as well as how to evaluate intent in the domestic violence setting. Obviously, the majority opinion will be much cited in future domestic violence forfeiture cases to evaluate the type of evidence justifying forfeiture:

Acts of domestic violence often are intended to dissuade a victim from resorting to outside help, and include conduct designed to prevent testimony to police officers or cooperation in criminal...
prosecutions. Where such an abusive relationship culminates in murder, the evidence may support a finding that the crime expressed the intent to isolate the victim and to stop her from reporting abuse to the authorities or cooperating with a criminal prosecution—rendering her prior statements admissible under the forfeiture doctrine. Earlier abuse, or threats of abuse, intended to dissuade the victim from resorting to outside help would be highly relevant to this inquiry, as would evidence of ongoing criminal proceedings at which the victim would have been expected to testify.\textsuperscript{88}

The majority separates its reference to the defendant’s dissuading the victim to obtain help from its mention of the existence of ongoing criminal proceedings. This disjunction appears to eliminate any requirement for an ongoing prosecution, generally implying that there need not be a current case pending to find forfeiture. However, the dicta is directed to abuse that culminates in murder, leaving open whether the majority would infer intent in other types of prosecutions, or whether felonies would be treated differently than misdemeanors.

The concurrence of Justices Souter and Ginsburg includes the following view of domestic violence evidence of forfeiture:

\begin{quote}
[T]he element of intention would normally be satisfied by the intent inferred on the part of the domestic abuser in the classic abusive relationship, which is meant to isolate the victim from outside help, including the aid of law enforcement and the judicial process. If the evidence for admissibility shows a continuing relationship of this sort, it would make no sense to suggest that the oppressing defendant miraculously abandoned the dynamics of abuse the instant before he killed his victim, say in a fit of anger.\textsuperscript{89}
\end{quote}

This language does not limit inferred intent to murder cases, and the rationale provided describing the dynamics of domestic violence extends to all manner of domestic violence prosecutions. The broad language suggests that proof of the abusive relationship is all that is needed for transferred intent. Indeed, the dissent interprets Justice Souter’s concurrence as meaning “that a showing of domestic abuse is sufficient to call into play the protection of the forfeiture rule in a trial for murder of the domestic abuse victim.”\textsuperscript{90} In contrast, the majority indicates such evidence “may” support a finding, inferring that some specific acts or statements of the defendant

\textsuperscript{88} Giles, 128 S. Ct. at 2693.
\textsuperscript{89} Id. at 2695 (Souter, J., joined by Ginsburg J., concurring).
\textsuperscript{90} Id. at 2708 (Breyer, J., joined by Stevens & Kennedy, J.J., dissenting).
are required to indicate an intent to isolate or dissuade the victim from obtaining help.\footnote{Id. at 2693 (majority opinion).}

The dissent of Justices Breyer, Stevens, and Kennedy would not require any witness-tampering intent at all for forfeiture.\footnote{Id. at 2707-08 (Breyer, J., joined by Stevens & Kennedy, JJ., dissenting).} Thus, it is clear that they will interpret this requirement flexibly. They specifically note:

\[\text{[e]ven the majority appears to recognize the problem with its “purpose” requirement, for it ends its opinion by creating a kind of presumption that will transform purpose into knowledge-based intent—at least where domestic violence is at issue; and that is the area where the problem is most likely to arise.}\footnote{Id. at 2708 (emphasis omitted).}

Justice Scalia rejects the dissent’s characterization that this is “nothing more than ‘knowledge-based intent,’”\footnote{Id. at 2693 (majority opinion).} but his disclaimer appears to be one of degree, rather than kind. Thus, all of the justices would accept a flexible view of inferred intent in domestic violence cases, but the nuances of how to establish such intent, and how broadly to apply it is open to disagreement. Professor Lininger has recently proposed presumptive bright-line rules to govern claims of domestic violence forfeiture and provide some consistency in application.\footnote{See Tom Lininger, The Sound of Silence: Holding Batterers Accountable for Silencing Their Victims, 87 Tex. L. Rev. 857, 892-93 (2009); see also Myrna S. Raeder, Being Heard After Giles: Comments on The Sound of Silence, 87 Tex. L. Rev. 105, 105 (2009) [hereinafter Raeder, After Giles].}

I have always been of the view that murder is different in the domestic violence context because the victim’s death is often accomplished in ways that are aimed at frustrating prosecution. Previous violence and threats instill fear in the victims who downplay their risk of continuing danger, and their murders are often accomplished at home without witnesses. Thus, prior to \textit{Giles}, I argued that previous statements of the victim should be viewed akin to dying declarations of individuals who are not isolated,\footnote{See, e.g., Raeder, After Davis, supra note 13, at 778-79.} relying on \textit{Mattox v. United States}.\footnote{156 U.S. 237 (1895).} \textit{Mattox} recognized that rules of law, “however beneficent in their operation and valuable to the accused, must occasionally give way to considerations of public policy and the necessities of the case.”\footnote{Id. at 243.} Ironically, while Justice
Scalia’s opinion adopts the rationale that abusive relationships may support inferred intent, the two concurrers and three dissenterers appear to form a different majority who adopt this rationale as proof of inferred intent without more. Thus, I expect the only forfeiture problem in murder cases will be when no classically abusive relationship can be established.

In contrast, I have been more hesitant to substitute evidence of an abusive relationship as evidence of forfeiture without more when the complainant is alive but refuses to testify, since so many complexities about the relationship confound an automatic finding that the defendant is the cause of her unavailability. In other words, that approach ignores reasons as to her unavailability that cannot be attributed to acts of the defendant.” Pre-Giles, I assumed specific evidence would be required, though “patterns of abuse and any posttraumatic stress disorder symptoms [w]ould be factored into the analysis.” Yet, in future cases, the dissenterers and concurrers could form a majority that would neither confine forfeiture to murder cases, nor require any evidence other than that of an abusive relationship. Paradoxically, grafting the modern view of the dynamics of domestic violence onto an originalist framework may eliminate not only the traditional requirement of any intent to witness tamper, but also the requirement of any specific showing of inferred intent beyond evidence of the abusive relationship in all cases, not just murder.

In addition, Justice Breyer’s suggestion that states may accept broader forfeiture views in a nontestimonial context has already borne fruit in Indiana, where Roberts v. State held that a defendant forfeited any objection under the rules of evidence. Obviously, evidence of the abusive relationship is much less costly for the prosecution to obtain than specific evidence of witness tampering, unless the victim is uncooperative from the outset. Even without cooperation, previous complaints by the victim to the police can be evaluated by the court, since forfeiture would be decided under

---

99 Raeder, Remember the Ladies, supra note 25, at 363-64.
100 Myrna S. Raeder, Confrontation Clause Analysis After Davis, 22 CRIM. JUST. 10, 19 (Spring 2007) [hereinafter Raeder, Confrontation Clause].
102 894 N.E.2d 1018, 1024-25 (Ind. Ct. App. 2008) (“[A] party who has rendered a witness unavailable for cross-examination through a criminal act, including homicide, may not object to the introduction of hearsay statements by the witness . . . under the Indiana Rules of Evidence.”).
F.R.E. 104(a), which permits consideration of hearsay. As I have argued elsewhere, the F.R.E. 104(a) standard permits the judge to consider character evidence including prior acts and expert testimony concerning the defendant’s abusive personality in determining the existence of forfeiture.

As I have also previously argued elsewhere, to the extent that an abusive relationship cannot be shown, in misdemeanors cases where the defendant is already on probation, parole, or supervised release, the better course may be simply to argue for the most severe penalty at revocation, since the Confrontation Clause is not implicated at such hearings, and the standard of proof is typically by a preponderance. Moreover, the absence of cross-examination satisfies due process when a sufficient explanation exists. The victim’s failure to cooperate or incompetency should supply good cause, and the statements would usually be found reliable because most of them would be admitted as excited utterances or through ad hoc exceptions that require trustworthiness. Ironically, while Hammon’s conviction was voided on Confrontation Clause grounds, he was also found guilty of a probation violation, and given his relatively short criminal sentence, the same penalty might have been reached via the revocation alone. The major difficulty with this approach is that it downplays the significance of the current crime, unless the defendant has previously committed a felony, which still offers the possibility of a substantial penalty.

CONCLUSION

Where has my foray into testimonialism and forfeiture taken me? Some may be surprised by my willingness to interpret testimonialism narrowly, since it results in a diminishment of cross-examination, something I have railed against for many years. However, I have always favored a balancing approach. And as this essay demonstrates, I have never been a fan of testimonialism because it disregards the core value of confrontation in relation to large quantities of

103 See Fed. R. Evid. 104(a).
104 Raeder, After Giles, supra note 95, at 113.
105 Raeder, Confrontation Clause, supra note 100, at 19.
106 See Black v. Romano, 471 U.S. 606, 612 (1985) (“[P]robationer is entitled [under due process] to cross-examine adverse witnesses, unless the hearing body specifically finds good cause for not allowing confrontation.”).
nontestimonial hearsay offered against the defendant, while at the same time dramatically impacting the prosecution for events outside the government’s control, even when the hearsay is reliable and critical to conviction.

Not surprisingly, courts have attempted to evade the testimonial ban by finding the admission of testimonial evidence to be harmless error. This has been particularly evident in cases where the claim of forfeiture is rejected. The Supreme Court recently noted, “[w]here a decision has ‘been questioned by Members of the Court in later decisions and [has] defied consistent application by the lower courts,’ these factors weigh in favor of reconsideration.” One wonders whether the Court will one day reject Crawford and its progeny like it previously rejected Ohio v. Roberts, to return to a more balanced Confrontation Clause approach as suggested by Mattox v. United States.

---


109 448 U.S. 56 (1980); see supra note 46 and accompanying text.

110 156 U.S. 237 (1895); see supra notes 97-98 and accompanying text.
Inquiry, Relevance, Rules of Exclusion, and Evidentiary Reform

D. Michael Risinger†

We are metaphorically assembled in this volume in celebration of the career of Margaret Berger, whose ideas on evidence and proof have enriched all of our work. Throughout her career, she has been concerned with both the theoretical and practical aspects of our enterprise as scholars. In such a setting, I hope that a reverie on fundamentals, followed by a modest call for reform, will not be looked upon as out of place.

It seems to me that there are some foundational issues in regard to the very notions of relevance and of rules of exclusion that can profitably be reexamined. First, I will assume that the issue of “exclusion” of information will arise in the context of an inquiry, that is, some human activity undertaken to at least attempt a determination about an issue for which the answer is initially unknown. Second, I will assume that the issue involved is often (but not always) properly characterized as an issue of fact in its most basic sense, that is, a question the answer to which will be a proposition about a specific empirically determinable state of the world (assumed to be) exterior to human consciousness. It is to such issues I will confine myself, at least initially.

In any situation fitting this description there will be a human inquirer who, at the beginning of the process, is in a poor state of knowledge, that is, the inquirer’s state of available

† John J. Gibbons Professor of Law, Seton Hall University School of Law. An early version of this paper was presented at the Miniforo on Legal Epistemology organized by Larry Laudan and held at the National University of Mexico in June, 2008. I would like to thank Professor Laudan for inviting these reflections, and the participants in the Miniforo on Legal Epistemology, particularly Ronald Allen and Erik Lillquist, for stimulating further reflections. I would also like to thank Craig Callen, Dale Nance, Roger Park, Charles Sullivan and Peter Tillers for very helpful comments on drafts of this paper, and Lesley C. Risinger for the usual indispensable aid, substantive and editorial.

† By using the word “determinable” I do not mean to commit myself to any strong claims about the perfection of knowledge, as will become obvious from the rest of the paper.
information is such that a finding² concerning the fact in issue would not be well-warranted. The point of the process is to enable the inquirer to have the information necessary either to make a properly warranted finding, or to conclude that such a finding would not be properly warranted based on the information obtained through the process.³

So far, we have not imposed any constraints upon the process involved, beyond whatever is implied by the word “process.” The process could be time-bound, or not. The process could involve a single inquirer or a group. The process could require the declaration of a conclusion, or not. The process could authorize the inquirers to obtain information (investigate) themselves, or disable inquirers from so doing, in whole or in part, after the process begins (thus splitting the decisional aspects of an inquiry from the investigative function). Such latter process would require or authorize investigation by non-decisionmakers, with an eye to potentially providing the results of such investigation to the decisionmaker, whether or not the decisionmaker could also investigate independently of others. It could allow some forms of investigation but not others, or rule out of bounds some forms of information.

This “ruling out of bounds” is the essence of an exclusionary rule. It is in the context of an inquiry involving a split between investigation and decision agents where we are most used to seeing exclusionary rules advanced or attacked, but the notion of an exclusionary rule is not theoretically limited to this context. Of course there might be practical constraints on inquirers making good faith efforts to follow mandated exclusions in a unitary function inquiry. Indeed, in the legal context, it is sometimes said that exclusionary rules in unitary systems make no sense because judges must look at the evidence to determine if they should not see it.⁴ But this goes too far. They need not look at the content under all forms of exclusionary rule. Sometimes they can merely look at the package. A rule directing a judge to “consider nothing in writing” would not necessarily expose a judge to any

---

² I use the word “finding” and “determination” interchangeably.
³ The issue concerning how much information is enough for a properly warranted finding in any given inquiry is a policy issue concerning the proper standard of proof, to use the legal term. In many non-legal settings each individual may be free to adopt whatever standard seems appropriate to them.
information beyond the fact that a writing exists, before the judge put the writing out of sight and out of consideration. Of course, most proposed exclusionary rules are not like this, but it is important to keep in mind that whether an inquirer is exposed to proscribed information in the process of determining its proscribed status falls along a continuum ranging from the necessity of full exposure to a situation where very little exposure is required, depending on the nature of the rule.

Which brings us to the question of what a “free proof” process of inquiry would look like. I take as my model of such a process an inquiry by an individual into whether, say, other humans had advanced knowledge of Lee Harvey Oswald’s intention to take his Mannlicher Carcano rifle to work and attempt a shot at John Kennedy from the 6th floor window of the Texas School Book Depository on November 22, 1963. From the perspective of this notional inquirer, the notion of pure free proof makes sense. The inquirer is totally free to consider whatever the inquirer believes to be information bearing on the inquiry, to take as much time as necessary to come to a conclusion, to come to no conclusion, or to any conclusion whatsoever, to any degree of certainty that results from the process and the

---

5 I am here using the term “free proof” in what is probably its most commonly assumed meaning in English, that is, in its strongest sense, the complete absence of any rules of exclusion, and in a weaker sense, the absence of all rules of exclusion except irrelevance. Deirdre Dwyer traces this use of the term to Thayer. See Deirdre M. Dwyer, What Does It Mean To Be Free? The Concept of “Free Proof” in the Western Legal Tradition, 3 International Commentary on Evidence 2 (2005), available at http://www.bepress.com/ice/vol3/iss1/art6/. The other dominant meaning of the term refers to the unconstrained freedom of the factfinder to evaluate what is in front of the factfinder in any way that seems best, unconstrained by any other arrangements such as rules of fixed evaluation (corroboration, “half-proof” etc.), presumptions, authoritative comments on inferential strength, whether mandated or simply authorized, etc. I would prefer to call the latter “free inference.” (And, in fact, this is better in line with post-revolutionary French terms that are the source concepts, liberté des preuves for the first and liberté d’appréciation for the second. See id. at 7.) “Free proof” in these terms thus concentrates in a legal context on the provision of information (or asserted information) and “free inference” on the use of whatever information is provided, so that free proof describes the freedom to proffer, and free inference the freedom to evaluate what is proffered in any way that seems appropriate, given the task to be undertaken by the factfinder (which in a legal setting would still be defined by the substantive law through instructions). A further step, which might be said to prevail in an extreme system of equity discretion, might be called “free decision” (and might encompass such things as jury nullification), but that is way beyond the scope of this paper.

6 For purposes of this little reflection, I am assuming (perhaps counterfactually) that there is no reasonable controversy that Oswald did take his rifle to work with an intent to take a shot at President Kennedy from the 6th floor of the Texas Schoolbook Depository building on November 22, 1963.
information discovered and considered. If the inquirer takes ouiji board pronouncements as authoritative (relevant, strongly probative) there’s no law (or a priori rule) against it.

If, on the other hand, the individual wishes to persuade others of any conclusion reached, the individual would be well-advised to take into account the epistemic investments of the target audience, because what the individual regards as providing a strong warrant for the conclusion may not be so taken by others. In that case, when the individual sets out to persuade, or to provide a warrant seen as acceptable to the audience, the audience may in practical terms impose a rule of exclusion on the one seeking to persuade: no reference to ouiji board results should be given, tendered, proffered, presented or suggested.

Well, you might say, that would not be a systemic rule, really, but a rule of persuasive or rhetorical prudence. And so it would be, if the context involved were merely history as a hobby, or a free intellectual pursuit. But now assume some official action, which will benefit some and be a cost to others, is made to turn on the results of the inquiry. Now there is more at stake than merely personal investments (for one thing, the newly imposed functional conditions require some time limits for decision). It would seem that now there must be an officially declared criterion of some sort concerning what is in bounds, and what is out of bounds, as providing material for an acceptable belief warrant. This circumstance is heightened in any such process (which by definition involves conflicting interests) in which representatives of the conflicting interests are allowed to proffer information to the inquirer. Real free proof would permit the destruction of functionally required time constraints by filibuster, and in many cases one side would have an incentive to undertake such a strategy.7

---

7 Otherwise known as a channeling board, spirit board, or talking board. I have adopted the popular spelling “ouiji” (also sometimes spelled “weegee”), which reflects the common pronunciation, in preference to the original spelling “Ouija,” which is trademarked.

8 A similar observation was made in that classic of modern legal relevance theory, George F. James, Relevancy, Probability and the Law, 29 CAL. L. REV. 689, 701 (1941). The James article is justly famous, being cited as a source for the approach of FRE 401 in its Advisory Committee note, FED. R. EVID. 401 advisory committee’s note, among other things, and at 17 pages is one of the shortest law review articles of lasting theoretical impact in the modern era. This is especially striking, considering the fact that James was predominately a tax/estates and trusts practitioner who was only in the academy (as an assistant professor at the University of Chicago) for less than five years in his mid-30s, and that he only wrote two articles on evidence topics (out of 11
Nor is it a simple solution to this problem to invoke the concept of relevance as the sole criterion to be applied to items claimed to be properly considered as “information” in such a context, because relevance is a much more problematic concept than most people realize. I think there are generally two contrasting ways to approach the notion of relevance, which I will call the “god perspective” and the “processor state perspective.” From the god perspective, there is no limit to the concept of relevance. Even if we conceive of a god-being who is not in some sense omniscient to begin with, the minimum assumption of the god perspective is that the god inference maker knows all that is necessary to make as much accurate inference as possible from any given item proposed for consideration, and has no time or processing capacity constraints. In such a circumstance, there is no obvious stopping point between any piece of factual information and all the facts that can be—the Laplacian hyper-inference grok. Everything becomes relevant to everything in a virtually tautological sense, or at least there is no knock-down reason to believe it doesn’t. So the god perspective on inference is, for
any pragmatic human use (including use as a limitation on filibuster), worthless, yet it is the stance that seems to be assumed in many legal rules and discussions of inference, particularly rules like Federal Rule of Evidence 401.

Another conceptual problem avoided by rejecting the god view is the problem of what has been called "unknown relevance," that is, evidence that is "relevant" but whose "probative value [is] entirely unknown, at least to an ordinary reasonable trier of fact." Dale Nance, Conditional Relevance Reinterpreted, 70 B.U. L. Rev. 447, 456 n.30 (1990); see also Richard Lempert, Modeling Relevance, 75 Mich. L. Rev. 1021, 1029-30 (1977). In my view, such proffers, while perhaps potentially relevant under other conditions, are simply "irrelevant" in any meaningful sense. This can be an especially problem when legislatures, for whatever reason, attempt to declare information categorically admissible when the covering generality connecting the evidence to the desired conclusion would be in many or most cases unknown to judge and jury alike, as in the case of the presumption of importation of marijuana from the mere fact that a seized substance was identified as marijuana, which was one subject of examination in United States v. Leary, 395 U.S. 6 (1969). See Harold A. Ashford and D. Michael Risinger, Assumptions, Presumptions, and Due Process in Criminal Cases: A Theoretical Overview, 79 Yale L.J. 165, 205-08 (1969).

This approach is sometimes referred to as the theory of "logical relevance" (though why it is thought to be more "logical" than other approaches is not entirely clear) and sometimes as the theory of "minimal relevance." Professor Callen attributes its initial explicit exposition to Professor George James, and the "minimal relevance" terminology to Professor Tillers. See Craig R. Callen, Rationality and Relevancy: Conditional Relevancy and Constrained Resources, 2003 Mich. St. L. Rev. 1243, 1254 n.51, 1280 n.182 (2003). The limitlessness of "minimal relevance," that is, relevance when viewed as a quality of the relationship between a probans and a probandum in the abstract, has been noted before, most specifically by Jerome Michael and Mortimer Adler: "If the criterion of admissibility were simply relevancy, in the strict logical sense . . . [then] nothing would be inadmissible since, as we have seen, nothing would be irrelevant in that sense." Jerome Michael & Mortimer Adler, The Trial of an Issue of Fact: II, 34 Colum. L. Rev. 1462, 1462 (1934). See the discussion in Peter Tillers, Reviser's Note, § 37.3, IA Wigmore on Evidence 1027-28 (Tillers rev. 1983) [hereinafter Tillers]. This limitlessness is mainly a product of two things: the first is the likelihood that rational inference is best formally modeled by a structure where a probans is related to a probandum by the "nomological glue" (to use Professor Tillers's term) of a covering generality of some sort. The second is the fact that the identity and extent of such covering generalities in the universe is currently unknown (and, given a variety of constraints flowing from the human condition, in principle unknowable), so that the relevance of any probans to any probandum can never be said with certainty not to exist as a function of a covering generality not currently known. The best illustration of this inescapable problem that I am aware of comes from Professor Callen (discussing it in the context of "conditional relevance"):

There are some situations in which our limited knowledge and cognitive resources make relevancy of evidence conditional on information about additional facts—regardless of whether we have a doctrine called conditional relevancy. Suppose that D is accused of committing homicide by shooting. No bullet has been recovered, nor did police find any trace of one at the scene of the crime other than the wound. Evidence is offered that D used a number of extremely cold substances in his laboratory research. On its face, that evidence would be irrelevant to most people. That is, it seems unlikely that there is any connection between murder by shooting, absence of a bullet, and experience with extremely cold substances such as liquid nitrogen. If, on the
The alternative notion of relevance takes into account not only the information being proffered as relevant to some other claimed fact, but also the characteristics of the processor that will mediate between the initial information and the asserted inference.  

Here we need to consider what we mean by “information.” Information is something that interacts with a decisionmaker (processor), broadly defined, which increases the rational warrant for some decision or group of decisions over potential rivals. Again, in approaching the concept of information in this way, I am emphasizing that the status of a stimulus as “information” is not inherent solely in the stimulus. It is dependent upon the way the stimulus interacts with the decisionmaker. Thus, whether a stimulus counts as information is a characterization of its interaction with and effect on a decisionmaker. No decisionmaker, no information, although things in the world that are stimuli that potentially could affect some decisionmaker under conditions not now prevailing might be called “potential” information. Not only does the status of something as information depend on its interaction with a decisionmaker, it must interact in a specified

other hand, there were reason to believe that (i) one could shape some such frozen gas to form a bullet, (ii) fire it from an air gun causing a fatal wound, after which (iii) it would sublimate, then that would make D’s access to such substances relevant. Admitting all proffered evidence on the theory that some fact making it relevant might turn up would be incredibly wasteful. Callen, supra, at 1278-79.

The Federal Rules of Evidence move all processor considerations to Rule 403, which I believe leaves almost nothing for Rule 401 to do on its own, at least if applied as drafted. I am not the first to make this observation. See Tillers, supra note 14, at § 37.2. (“[T]here are very few cases in which the exclusion of evidence can be explained on the ground of irrelevance alone. This ironic result is the product of the constricted exclusionary force of the legal principle of relevancy that arises from the assumption that the only legally irrelevant evidence is that evidence that also happens to be ‘logically’ irrelevant.” Id. at 1021.)

This definition of information proper, as distinguished from “potential information,” was first suggested, sub nom. “actual relevance” and “potential relevance” in D. Michael Risinger, John Henry Wigmore, Johnny Lynn Old Chief, and “Legitimate Moral Force,” Keeping the Courtroom Safe for Heartstrings and Gore, 49 Hastings L. J. 403, 431-35 (1998). It differs from the definition of “information” used in communication theory, where information is defined so broadly as to include both what is here referred to as information proper, and also what is referred to as “potential information.” See Warren Weaver, “Recent Contributions to the Mathematical Theory of Communication,” in Claude Shannon & Warren Weaver, The Mathematical Theory of Communication (1949): “The word information in this theory is used in a special sense that must not be confused with its ordinary usage. In particular information must not be confused with meaning.” Id. at 8.
way. To count as information, it must both affect a decision, and affect it in an accuracy-improving way.

By accuracy-improving, I mean that the stimulus must reliably be the kind of stimulus that improves the likelihood of the correspondence of a decision with the characteristics of the exterior world (again, we are limiting ourselves to decisions about facts, or to the fact component of more complex fact-value decisions). By defining information in this way I realize that I am making some fundamental commitments: first, to some version of a theory of knowledge that has both probabilistic and correspondence aspects; and second, to the proposition that whether a stimulus is information depends both, and as much, on the state of the decisionmaker as on the state of the stimulus. I will not attempt a full-scale defense of the first here, since I think that such an approach to knowledge is inherent in most contexts (for instance, legal contexts) to which I will apply the approach. The second requires both more unpacking of its implications, and more defense.

One of the implications of a Laplacian determinist account would be that, from the perspective of a being that knew all the details of what was taken (by Laplace) to be a fully specified formal system (and perhaps more), there is no information in the sense I have defined it. As previously noted, every stimulus entails immediate complete knowledge of all other details of the system at all possible times past and present. I will not burden the reader with a rehearsal of the many problems of adopting this view as one that defines the nature of some ultimate physical reality.¹⁷ But the mental experiment entailed in considering this view results in one important point. “Decisions” about states of the world can only exist and be “informed” by “information” exactly because we live in a world of imperfect knowledge where virtually everything must be approached based on a probabilistic evaluation of the interaction of stimuli we receive acting upon the current states of our neural processors (that is, our reactions to those stimuli, conscious, semi-conscious or unconscious).

Something that is information in regard to a particular issue is relevant to that issue. Something that is not information is not relevant. If something is information it is therefore evidence in the core sense, that is, it will affect the

¹⁷ See supra note 9.
processor (decisionmaker) in a way that is a net benefit to probabilistically better results, based on a pertinent evaluation of the information (about which more later). Therefore, whether something is relevant to a decision can only be determined by knowledge or assumption concerning the characteristics of that which is proffered (something claimed to be “evidence,” i.e. relevant information) and the characteristics of the decisionmaker, and the interrelationship between the two.

Consider this: a dog is caught on a ledge. He is skittish. He may perform an act injurious to himself. Does this dog make a decision when he moves? Assuming the answer is in some sense yes, what stimulus could I give him that would count as information relevant to his decision? Words in English explaining the effects of falls from heights, or the value of remaining calm, or the safest route off the ledge? If not, then this illustrates the difference between potentially and usably relevant information, and how the two are a function of both the stimulus and the state of the putative decisionmaker.

So something is usably (as opposed to potentially) relevant information to a decision (that is, only then can it be counted as information at all) only in regard to the characteristics of the processor—is it the kind of information that raises the likelihood of decisions that conform to exterior empirical reality? I am not here making any claim about what the characteristics of an ideal processor might be. The

---

18 Or, one might say more accurately, between stimuli that are candidates to be considered information under different conditions than those currently prevailing, and those that actually are information.

19 But if I hold out a steak to lure the dog off the ledge, have I in any sense given him information that informs his decision about the problem of the ledge? Whatever drives are behind his action, and however we might move away from purely behavioral accounts to hypothesize an internal processing that might count as a decision about the steak, his response would not seem, at least at first blush, to represent a solution to the problem of deciding about ledge choices (except perhaps by assuming some connection between approaching humans holding steak and something counting as general trust).

20 Well, only very weak claims. It would seem that a stimulus that caused a response that induced an action for reasons unrelated to the pending problem, as might be the case in the dog example in note 19 supra, would not count as information even though it caused something that might be labeled a decision that resolved the problem. The steak resolved the problem, but not by providing an answer to an inquiry. Beyond that, as Professor Tillers has warned us for years, it is incumbent on us not to confuse our (current) models of ideal “rational” decision (which in themselves may turn out to be less than ideal) with the way humans process information or make optimal decisions about facts. See TILLERS, supra note 14, at 1017; see also Peter Tillers, Are There Universal Principles or Forms of Evidential Inference? Of Inference Networks and Onto-Epistemology, in JOHN JACKSON, MAXIMO LANGER & PETER TILLERS, EDS., CRIME, PROCEDURE AND EVIDENCE IN A COMPARATIVE AND INTERNATIONAL CONTEXT: ESSAYS
processor that counts is the processor we are ultimately stuck with, that is, humans en grosse. So if humans do not conform to an ideal processor, being just what they are and representing the range that they do in terms of the range of neural processor space and speed, prior experiences and heuristics, etc., then rules of admission (and a fortiori, rules of exclusion) must be fashioned with the best available information about these characteristics in mind, and more such information should be developed and considered for the improvement of that process of decision.

Nothing I have said thus far is necessarily particularly novel, of course. It is consistent with general notions

IN HONOR OF MIRJAN DAMASKA (2008). This is not to say that anything goes, or that human mental processors can be neither criticized nor trained to be more accurate in evaluation, but only that our natural modes, appropriately disciplined, may have epistemic advantages over currently available theoretical models. See Callen, supra note 14, at 1258-78. It seems to me that Daniel Dennett is right to say that whatever thinking (especially inferential thinking) is, it is virtually certain to be some sort of computational process—not one mimicking our current notions of ideal computation perhaps, but rather one that maximizes the trade-off between computational accuracy and efficiency through the use of modules, heuristics, and hierarchical selection of preliminary results. See generally DANIEL C. DENNETT, CONSCIOUSNESS EXPLAINED (1991), particularly ch. 9, “The Architecture of the Human Mind,” at 253-82. It seems also very likely that some of its heuristics are hard-wired, some come more-or-less default wired subject to revision, while others really are the product of experience repeated and internalized. These processes provide the generalities necessary to make sense out of case-specific evidence, see WILLIAM TWining, RETHINKING EVIDENCE: EXPLORATORY ESSAYS 332-40 (2d ed. 2006), and create the implied reference classes that make the problem of relevance not a purely logical exercise even after the elimination of the god view. See generally Ronald J. Allen & Michael S. Pardo, The Problem of Mathematical Models of Evidence, 36 J. LEGAL STUD. 107 (2007).

It also seems reasonably clear that some of our processing structures, heuristic, analogical and story-influenced as they appear to be, work well under some conditions and less well under others, leading us predictably astray, for instance, in attempting to play three-card monte. See D. Michael Risinger & Jeffrey L. Loop, Three Card Monte, Monty Hall, Modus Operandi and “Offender Profiling”: Some Lessons of Modern Cognitive Science for the Law of Evidence, 24 CARDOZO L. REV. 193, 193-210 (2002). I am inclined to evolutionary explanations for this state of affairs, but this is controversial and unnecessary to the present paper. At any rate, our evidence rules and practices should attempt to understand these conditions and take them into account.

Indeed, it resonates closely with many of the points made (in the context of a discussion of “conditional relevancy”) in Craig Callen’s fine article, Conditional Relevancy and Constrained Resources, supra note 14. Consider, also, the following from Professor Allen, discussing the limitations of a pure Bayesian account of proof at trial:

Whereas under the unconditional probability assumption there are too many possible accounts of reality, under the conditioned-on-trial-evidence assumption there are too few—none actually. There are none because this possibility suffers from an infinite regress of a different sort from the unconditional probability assumption. The regress here comes from the fact that evidence does not announce its own implications, those implications emerge from the effort of human contemplation.
concerning the implications of naturalized epistemology in the Goldman vein applied to evidence law to be found, most specifically, in Ronald Allen and Brian Leiter’s important article in the Virginia Law Review in 2001, though I was also moving in something of a similar direction myself somewhat earlier.

Allen and Leiter adopt the position that proper rulemaking in a jury system requires the rulemaker to take into account both “the epistemic frailties of jurors, and the epistemic limits of rule-appliers [the “gatekeepers”], namely, judges.” And later in the article they assert that the main question to be asked in regard to any rule of inclusion or exclusion is “an essentially empirical question: Does this rule of inclusion or exclusion in fact increase the likelihood that factfinders, given what they are actually like, will achieve knowledge about disputed matters of fact?” They also appear to adopt the primacy of information derived from formal empirical studies as the main appropriate source of information about those epistemic frailties, rejecting “rootless theorizing” as too “a priori.” However, it seems to me that their approach to the results of such studies, and to other sources of relevant knowledge on the advisability of rules and

Ronald J. Allen, Factual Ambiguity and a Theory of Evidence, 88 NW. U. L. REV. 604, 613-14 (1994). This emphasis on the “effort of human contemplation,” developed at various later points in the article, see, e.g., id. at 619-20, appears to be the same as an emphasis, for purposes of relevance analysis, on the state of the humans doing the contemplation, i.e., the factfinder(s).

See generally ALVIN I. GOLDMAN, EPISTEMOLOGY AND COGNITION (1986).


See D. Michael Risinger, supra note 16, at 403, 431-46 (“The Relevance of the Irrelevant”) (drawing the distinction between potential and usable relevance):

[Federal Rule of Evidence 401] declares evidence relevant if it has “any tendency to make the existence of a fact more or less probable than it would be without the evidence” without suggesting a referent to the trier of fact’s rational capacities to derive or process the information. It emphasizes the content of the code independent of the characteristics of the decoder.

Id. at 433 n.79; see also D. Michael Risinger, Preliminary Thoughts on a Functional Taxonomy of Expertise for the Post-Kumho World, 31 SETON HALL L. REV. 508, 518-20 (2000).


Id. at 1537.

Id. at 1521-26.
rule changes, is so cautious that it just about rules out a “naturalized” reform agenda.

The main tool Allen and Leiter use to arrive at this position is the “external validity” play, although they don’t use that label. Whenever formal data are derived from experimental or quasi-experimental studies, even of the best designed sorts, there is always an issue of how far the results can be generalized to other universes of phenomena beyond the exact set studied.28 This question is perhaps trivial in physics, because of well-warranted assumptions of fungibility, but as such fungibility assumptions become less and less tenable, external validity concerns rise, and when we get to issues of human behavior, these concerns are at their highest. Allen and Leiter rightly recognize that there is always a question in generalizing from behavior under simulated test conditions to behavior under real world conditions, and that this may be especially true when attempting to generalize to behavior under the very unusual conditions presented by the context of a jury trial. This can be seen as laudable cautious skepticism, and I must say that as to each individual proposition considered, I can view it that way myself. However, globally, I can’t escape the feeling that they have set the bar too high. In regard to the reform implications of the demeanor studies, they accept the conclusions of Professor Wellborn that these studies do not yet compel any specific changes in the current way we do business.29 They then turn to two other areas, probabilistic evidence and character evidence, which they say “cry out for reform and/or additional research.”30 But in the end, after looking at the condition of the empirical record and examining it through their external validity lens, no suggestions for possible reform are forthcoming.31

Now I concede that I am probably being a bit cranky about this. The line between proper circumspection about

28 The *locus classicus* for the consideration of the problem (and the source of the term) is the work of Donald T. Campbell. The most recent version of this work is to be found in WILLIAM R. SHADISH, THOMAS D. COOK & DONALD T. CAMPBELL, EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS FOR GENERALIZED CAUSAL INFERENCE 83-93 (2002). Reflections on external validity issues in legal contexts can be found in Mark P. Denbeaux & D. Michael Risinger, Kumho Tire and Expert Reliability: How the Question You Ask Gives the Answer You Get, 34 SETON HALL L. REV. 15, 25-28 (2003) and D. Michael Risinger, Innocents Convicted: An Empirically Justified Factual Wrongful Conviction Rate, 97 J. CRIM. L. & CRIMINOLOGY 761, 782-88 (2007).
29 Allen & Leiter, supra note 23 at 1540-42.
30 Id. at 1542.
31 Id. at 1545, 1549.
external validity and unreasonable status quo conservatism is difficult and hardly self-defining. But I come to this writing somewhat bent by watching the likes of Professor Ebbe Ebbesen play what I consider to be a radically skeptical version of the external validity card to aid various law enforcement status quo conservatives in resisting needed eyewitness identification reforms.32

In my view, what is needed for proper “naturalized reform” of the law is a combination of formal empirical data and critical common sense which can together form the basis for wise choices. Common sense in this context refers both to the sense of the world that humans obtain through the process of lived experience, and to processes of critical evaluation available in ordinary reflection and discourse that can be brought to bear on both the products of the laboratory and the less formal products of the laboratory of life.” It is commonly

32 Professor Ebbesen is a social psychologist who teaches at the University of California at San Diego. For a long time he has been one of the main prosecution witnesses called in litigation across the country to resist the admissibility of testimony of eyewitness researchers on the weaknesses of eyewitness identification testimony under various conditions. He has also been the go-to guy for providing academic support for law enforcement resistance to changing the way that eyewitness identification procedures are conducted and presented in court. See, e.g., United States v. Hines, 55 F. Supp. 2d 62 (D. Mass. 1999); People v. Smith, 784 N.Y.S.2d 923 (N.Y. Sup. Ct. 2004); People v. LeGrand, 747 N.Y.S.2d 733 (N.Y. Sup. Ct. 2002). He seems to have been the main design consultant (if it can be called a design) for the Illinois eyewitness study, the results of which were the basis of the infamous Mecklenberg Report. For a full examination of the many weaknesses of the Mecklenberg Report, see generally Daniel Schacter et al., Policy Forum: Studying Eyewitness Identifications in the Field, 32 LAW & HUM. BEHAV. 3 (2007). This evaluation was the product of a panel assembled by John Jay College of Criminal Justice. The members of the panel were Daniel Schacter, Professor of Psychology, Harvard University; Robyn Dawes, Queenan Distinguished University Professor, Carnegie Mellon University, Fellow, American Statistical Association; Henry L. Roediger III, James S. McDonnell Distinguished University Professor at Washington University, former President, Association for Psychological Science; Larry L. Jacoby, Professor at Washington University; Daniel Kahneman, Professor of Psychology, Princeton University, 2002 Nobel Laureate in Economics; Richard Lempert, Distinguished Professor, University of Michigan School of Law, and Division Director for the Social and Economic Sciences at the National Science Foundation, 2002-2006; Robert Rosenthal, Distinguished Professor at University of California, Riverside, and Pierce Professor of Psychology emeritus, Harvard University, Co-Chair Task Force on Statistical Inference, American Psychological Association. For another evaluation of the Mecklenberg report, together with a description of the kind of reform proposals regularly opposed by Prof. Ebbesen, see generally Richard A. Wise, Kirsten A. Dauphinais, and Martin A. Safer, A Tripartite Solution to Eyewitness Error, 97 J. CRIM. L. & CRIMINOLOGY 807 (2007).

33 What I am embracing is a version of what C.S. Peirce referred to as “critical common-sensism.” See Christopher Hookway, Critical Common-Sensism and Skepticism, 24 NOUS 397 (1990). I am aware of the sizable literature on the notion of common sense and its characteristics, strengths and weaknesses, although I must admit that I have only scratched its surface. A good set of references, although concededly now perhaps a bit dated, appears in Barry Smith, Formal Ontology,
observed that common-sense inquiry and scientific inquiry are not different in kind, and so it would hardly be surprising if a program of naturalization would supplement formal data with common-sense evaluation, since for most issues of interest for purposes of the law, definitive formal data either cannot be developed because of limitations on human research, or else they must await programs of research stretching into the far reaches of the future. Unless we supplement them with common sense, we seem doomed to embrace the status quo indefinitely.

In embracing a common-sense component to a program of naturalization, I am of course not really disagreeing with either Professor Allen or Professor Leiter, since they have both had positive things to say in the past about the role of common sense in things legal and evidential. So where might a

---


34 See generally Susan Haack, Defending Science—Within Reason (2003), especially ch. 4, “The Long Arm of Common Sense.”

35 Consider this from Prof. Leiter:

I find a certain type of pragmatism attractive—indeed, unavoidable—but it is both more modest and more radical than the apology for fuzziness that masquerades as pragmatism in the law journals. This pragmatism is a relative of the type one finds in philosophers like Carnap and Quine, and that has entered the philosophical lexicon in the metaphor of “Neurath’s boat.” The radicalism of this pragmatism resides in its recognition that the only possible criteria for the acceptance of epistemic norms—norms about what to believe—are pragmatic: we must simply accept the epistemic norms that work for us—that help us predict sensory experience, that allow us to manipulate and control the environment successfully, that enable us to “cope.” Pragmatic criteria are, at the limit, the only possible criteria for the acceptance of epistemic norms precisely because we can’t defend our choice of any particular epistemic norm on epistemic grounds ad infinitum. At some point, we must reach an epistemic norm for which the best we can say is, “it works.”

But which norms actually work for us? Take an example: “Don’t believe in a hypothesis that figures in a non-consilient explanation of experience” is a norm for belief—call it the “consilience” norm. A non-consilient explanation is one that posits an explanans—the thing that does the explaining—that seems too narrowly tailored to the explanandum—the event to be explained. Here’s how this consilience norm works in our lives. Suppose while sitting at home, all the lights in the house suddenly go out at the very same moment. What fact about the world explains this?

Explanatory hypothesis number 1:

Conspiring leprechauns have simultaneously thrown all the light switches in the house.

By contrast, explanatory hypothesis number 2 proposes that:
combination of formal data and common sense insight take us? Well, for one thing, I think that it would carry us toward supporting most of the reforms in eyewitness procedures that have been suggested by Gary Wells and others, perhaps along something like the following lines:

1. There are a larger number of erroneous convictions of factually innocent defendants than many believed or were willing to admit until recently. (This is supported both by data and by an examination of the statements concerning the rarity

There has been a general power failure, i.e., electrical current has stopped entering the house.

Both explanatory hypotheses suppose an ontology: mischievous leprechauns on the one hand; electricity, wires, and currents on the other. But the appeal to leprechauns is non-consilient: it seems a gratuitous ontological posit, precisely because supposing that leprechauns exist doesn’t help explain anything else. Their existence doesn’t explain our observations—we haven’t seen any—nor does it help explain the restoration of power—we neither need to “exterminate” the leprechauns in order to retain power, nor do we even need to turn on all the light switches they are hypothesized to have flipped. By contrast, assuming the existence of electrical currents proves a very fruitful ontological posit: it not only cues us to the appropriate steps to take to restore power in the house, but it helps explain a range of ordinary phenomena, like why the television goes off when unplugged from the socket. Since the consilience norm favors the electricity ontology over the leprechaun ontology, and since the former works better than the latter, it appears that a good reason to accept the consilience norm is because of its practical cash-value.

Indeed, the consilience norm—and its other relatives in a scientific epistemology—have worked very well for us humans: they helped depopulate our ontology of leprechauns and gods and ethers, and they are foundational norms in scientific practice, a practice that sends the planes into the sky, keeps the food from spoiling in the refrigerator, and alleviates human suffering through modern medicine. From a philosophical standpoint, what bears special notice is that the epistemic norms of common sense and the epistemic norms of science are simply on a continuum.


And this from Professor Allen:

There is plenty of work to be done figuring out how people reason, and particularly how they reason about legal affairs, even if we do not pursue these matters with the equivalent of highly general, top-down scientific theories or the tools of postmodern French literary theory. However, even if the fun quotient does go down, the significance quotient may go up. As the sociologist Lindenberg said: “Common sense finds its way into a body of law if it has a strong influence on social relations; central to all of these is the body of law governing evidence . . . . The body of law governing evidence may be the strongest bastion against sudden assaults on common sense.” I would add that resisting sudden assaults on common sense may be one of the most important guarantors of the continuing progression of civilization.

of such events over time, which I guess is another form of
data.36)

2. Laboratory data suggests that humans are
vulnerable to a variety of influences that lead to erroneous
selection of innocent persons as the perpetrators of crime in a
non-trivial number of cases, generally involving stranger-on-
stranger identifications.37 This result is not particularly
surprising from a critical common-sense perspective.

3. Data from other studies indicate that humans tend
to over-value eyewitness identification,38 especially if
confidently given, even though confidence is not necessarily a
good predictor of accuracy under many conditions.39 This result
also does not appear to be surprising from a critical common
sense perspective.

4. Data from other sources (examination of the records
of DNA exoneration) indicates that erroneous eyewitness
identification is the single most common factor involved in such
cases.40 This is also not particularly surprising.

5. Data from laboratory studies point to ways of
conducting criminal investigations involving eyewitness

36 See generally Risinger, supra note 28, at 765-68 et passim.
37 The literature on eyewitness identification research and its results is vast.
A good starting point is Gary L. Wells, Eyewitness Identification: Systemic Reforms,
38 See generally R.C.L. Lindsay, Expectations of Eyewitness Performance:
Jurors' Verdicts Do Not Follow from Their Beliefs, in ROSS ET AL., ADULT EYEWITNESS
TESTIMONY (1994).
39 Id.; see also, e.g., Siegfried L. Sporer, Choosing, Confidence and Accuracy:
A Meta-Analysis of the Confidence-Accuracy Relation in Eyewitness Studies, 118
40 BARRY SCHECK, PETER NEUFELD & JIM DWYER, ACTUAL INNOCENCE (2000)
(app. 2, tbl. 2, providing an analysis of 62 DNA exoneration cases showing inaccurate
eyewitness identifications in 84% of the cases); Michael J. Saks & Jonathan J. Koehler,
The Coming Paradigm Shift in Forensic Identification Science, 309 SCIENCE 892 (2005)
(analysis of 86 DNA exoneration cases showing inaccurate eyewitness identifications in
71% of the cases); Brandon L. Garrett, Judging Innocence, 108 COLUM. L. REV. 55, 78
(2008) (analysis of 200 DNA exoneration showing inaccurate eyewitness
identifications in 79% of the cases). The exact implications of these statistics are not
clear and must be approached with caution. As Roger Park has pointed out, in a set of
cases proved by DNA to be factually wrong, and made up largely of stranger rape cases,
the statistic could hardly be otherwise, and taken by itself is weak evidence that
eyewitness identifications are especially unreliable. See Roger C. Park, Eyewitness
Identification: Expert Witnesses Are Not the Only Solution, 2 LAW, PROBABILITY & RISK
305, 305-06 (2003). This is of course another variation of the “denominator problem.”
Without a reference class in which both the number of accurate and inaccurate
eyewitness identifications is known, it is not possible to derive a rate of inaccuracy.
However, the DNA data are hardly inconsistent with the general claim that eyewitness
identifications are surprisingly unreliable, a claim that predates DNA exoneration
significantly. See WILLIAM JAMES, THE PRINCIPLES OF PSYCHOLOGY 516 (1890), Great
identifications that significantly reduce the likelihood of erroneous identifications of innocent suspects.\textsuperscript{41} Some of these suggestions are, when examined in a critical common sense fashion, virtually cost free both as to monetary costs and costs in lost identifications of the guilty that are at all epistemically defensible to begin with. In such circumstances, such reforms should be undertaken.\textsuperscript{42}

Let me expand on the latter point. What I am referring to is the adoption of a requirement that all pre-trial identification procedures (photo-spreads and corporeal line-ups) be administered by someone who does not know which person in the array is the suspect and which is a filler, that is, the so-called “double blind”\textsuperscript{43} administration of the identification procedure. While it is possible to argue that other proposed reforms (such as sequential viewing to reduce selection by “relative judgment”) may have a cost in reducing some epistemically warranted accurate identifications, such an effect is not possible in regard to the masking of administrators. No one has yet come up with an argument that plausibly suggests how it could be that identifications which result solely from the administrator cueing, whether conscious or unconscious, and which would not be made by the witness independent of this variable, have any epistemic justification at

\textsuperscript{41} See Wells, supra note 37.

\textsuperscript{42} The presentation in the text should not be read to mean that I do not think that other aspects of the reform proposals should not be adopted. I selected the double blind aspect as the easiest to use as a clear illustration. Beyond that, for instance, I believe that a strong case can be made for adopting sequential presentation, even though the loss of selections which results from a sequential presentation can be more easily argued to represent the loss of a certain percentage of epistemically justified “accurate” identifications of true perpetrators. The set of identifications lost at the margin between simultaneous presentation procedures and sequential procedures is not large, and is likely to be rich in non-perpetrator selections, or false positives. The ratio of lost true positives to lost false identifications seems likely to be small enough to count as an acceptable cost for the reduction in innocents convicted under virtually any justifiable approach to such a “reform ratio.” See the discussion in Risinger, supra note 28, at 796-97.

\textsuperscript{43} This terminology is adopted by virtue of an analogy to “double blind” study design in various research contexts. In a double blind study, the test subjects do not know if the “treatment” they are subjected to is the actual test variable or a placebo (single blind) and the people interacting with the test subjects in the administration of the test do not know either (double blind). The term “double blind” in the eyewitness context is a bit out of kilter, since the notion of the original blind (the fact that the witness does not know specifically which person is the actual suspect) is entailed in the notion of a line-up style procedure (whether photo or corporeal) to begin with, so it seems in some ways that the term “blind administration” would be more natural in capturing the proposed reform, but “double blind administration” has become fairly standard in the literature.
all. To so argue is akin to arguing that we would be justified in retaining and not reforming a coin-flipping process in making selections from line-ups, because otherwise we would lose the one in X number of identifications that was coincidentally accurate. And all logistical and cost objections to blind administration have easy technological responses that render the objections trivial at best. So a combination of formal data about both the human circumstances that generate misleading evidence for use and trial, and juror inability to properly discount such misleading evidence once it reaches them, together with common sense evaluation of objections to reform, should lead to both mandated pre-trial procedures, and rules of exclusion to enforce the use of those procedures, all based on “naturalized reform” principles.

I could go on giving examples, but I won’t. In any case, I hope to have gone some way toward establishing that only by utilizing both formal data and critical common sense can a naturalized approach to rules of exclusion (and the proof process more generally) be of use in improving the product of adjudication.

44 See Letter of John J. Farmer, Jr., Attorney General of New Jersey, transmitting new statewide line-up identification guidelines to all New Jersey law enforcement agencies 2 (April 18, 2001), http://www.state.nj.us/lps/dcj/agguide/photoid.pdf (requiring blind administration where practicable, and recommending use of “[t]echnological tools, such as computer programs that can run photo lineups and record witness identification independent of the presence of an investigator”).

45 A reader might object that I had gotten away from the specific subject of exclusionary rules by using an example of reform up-stream from the trial. However, I have intentionally done this, for a number of reasons. First, in my opinion, if partisan adversary presentation and argument at the trial can be argued to add epistemic strength to the results of our trial system, the major epistemic weaknesses of our current adversary arrangements are the result of party control of the investigation and development of information for trial. Party interests being what they are, every opportunity for advantageous selection, distortion and massage is likely to be taken. The most pressing needs for reform are at this stage, to insure that information is both complete and as undistorted and uncorrupted as possible. Once we determine what processes should be mandated to this end, then rules of exclusion at trial (evidence rules stricti juris) must then be put in place to protect the requirement of the mandated pre-trial processes—proper manifestations of the best kind of best evidence principle embraced by Professor Nance. See Dale A. Nance, The Best Evidence Principle, 73 IOWA L. REV. 227 (1988) for Professor Nance’s classic revitalization of the “best evidence” concept.
Applying *Daubert* Inconsistently?

PROOF OF INDIVIDUAL CAUSATION IN TOXIC TORT AND FORENSIC CASES

*Joseph Sanders*

I. **INTRODUCTION**

*Daubert v. Merrell Dow Pharmaceuticals, Inc.* ushered in a new era in the assessment of expert testimony. *Daubert* and its two companion cases in the “*Daubert* trilogy,” *General Electric v. Joiner* and *Kumho Tire v. Carmichael,* drastically altered the law governing the admissibility of expert evidence. In the federal courts and in a substantial majority of state courts, the old *Frye* rule was swept aside for a new, multifaceted test. All three of these cases involved a causal

---

1 Thanks to Edward Cheng for helpful comments on an earlier draft. I am particularly honored to be part of a festschrift for Margaret Berger. One anecdote tells all. Not long after the Supreme Court's decision in *Daubert v. Merrell Dow Pharmaceuticals,* Margaret and I each published an article in the Minnesota Law Review. Our articles were grouped together as a "special issue" on *Daubert.* One of my colleagues later jokingly commented that I must be special to be one half of an entire special issue. I replied, no, Margaret's article was the special issue and I tagged along. By the way, her article, *Procedural Paradigms for Applying the Daubert Test,* 78 MINN. L. REV. 1345(1994) was an early beacon for judges grappling with this new admissibility rule.

2 *509 U.S. 579 (1993).*

3 *522 U.S. 136 (1997).*

4 *526 U.S. 137 (1999).*

5 The Federal Rules of Evidence have been changed to reflect these cases. Most importantly, Rule 702 now reads:

> If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.

FED. R. EVID. 702 (changes in italics).

6 *Frye v. United States,* 293 F. 1013 (D.C. Cir. 1923). According to the *Frye* test, scientific evidence should be admitted only when the scientific principle upon which the expert's testimony is based is “sufficiently established to have gained general acceptance in the particular field in which it belongs.” *Id.* at 1014.
question and challenges to causal assertions have remained at the very center of admissibility law. The costs and benefits of this filter on expert causal assertions has been the source of considerable controversy.6 One aspect of the controversy is a concern that the standard is not applied in a consistent manner.

Inconsistency appears in many guises. There is inter-jurisdictional inconsistency between the federal and state courts that have adopted Daubert and the state courts that continue to apply the Frye test.7 Even among jurisdictions employing the same test, there is substantial variation.8

More worrisome, perhaps, are inconsistencies within jurisdictions. In this regard, the most frequently discussed inconsistency is between civil and criminal cases. A number of people note that courts are more likely to permit causation experts, especially the state’s experts, to testify in criminal cases than in civil cases. For example, Professor Berger notes that,

In civil cases, courts engage in rigorous gatekeeping and often exclude plaintiffs’ experts because the theory underlying their testimony has not been adequately validated. But I see no sign of a parallel approach in criminal cases even [where] there are problems with the assumptions on which the prosecution's expert testimony rests.9

Others have made similar observations10 and this conclusion is supported by several empirical studies.11 On the other hand,
many have argued that the admissibility bar is set too high on the civil side.\footnote{12}

In this article, I argue that we misunderstand the nature and cause of the inconsistency when we lump together all toxic tort cases and compare them to all forensic cases. If we disaggregate the toxic tort admissibility opinions that deal with general questions (e.g., whether asbestos causes lung cancer), and those that deal with causal questions that relate to the individual plaintiff, we see that experts testifying on the latter causal question are judged by admissibility standards that are nearly as liberal as the standards applied in forensic cases. On the other hand, with respect to one type of forensic proof, DNA testimony, the courts impose an admissibility standard at least as high as that used for general causation cases in toxic torts.

It is the thesis of this article that the liberal standards applied with respect to specific causation and forensic experts have a similar source. They are the result of a gulf between the needs of the law and the products of science and they reflect a judiciary grappling with—or, in the forensic context, sometimes refusing to grapple with—the difficulties this presents for expert witness admissibility standards.\footnote{13}


\footnotetext[13]{After I had completed this article, David Faigman shared with me his contribution to this festschrift. His article, \textit{Evidentiary Incommensurability: A Preliminary Exploration of the Problem of Reasoning from General Scientific Data to Individualized Legal Decision Making}, 75 BROOK. L. REV. 1115 (2010), makes many of the points I make in this article concerning the gulf between the causal generalizations of most scientific inquiry and the specific causal analyses required in individual trials. See also DAVID L. FAIGMAN, \textit{LEGAL ALCHEMY: THE USE AND MISUSE OF SCIENCE IN THE LAW} 69 (1999).}
In Part II, I describe the gulf between the search for generalities in science and the need for particulars in law. In Parts III and IV, I discuss the effect of this gulf, first with respect to toxic tort “specific causation” experts and, second, with respect to forensic experts. In both areas, the relative weakness of the available empirical evidence leads courts to adopt liberal admissibility standards. Part V briefly summarizes the discussion in the previous two sections. Part VI argues that the liberal admissibility standards applied in these two areas are but two instantiations of law’s general contextual approach to knowledge.14 By and large, this contextual approach serves the legal system well, but in some situations it produces less than optimal levels of expertise. This occurs because the courts fail to adopt a contextual approach that attends to the future as well as to the case being decided. I outline the circumstances in which a longer view may serve us well and argue that those circumstances exist in the forensic arena. Part VII discusses alternatives open to the courts if they wish to adopt this longer view. The article ends with a brief summary of the argument.

II. THE SCIENCE-LAW DISCONNECT

There is a disconnect between science and law, and this disconnect helps to explain how law approaches certain types of causal questions. The disconnect is simply this: the law’s search for causal information about a particular case often finds little or no help from science. In order to understand this problem, I need to say a bit about the scientific enterprise. Those who study the doing of science would generally agree that there is no special “scientific method” that is different from and better than other ways of understanding the world.15 However, science is chock full of specific theories and methodological prescriptions concerning how to test these

---

14 A contextual approach varies the justification needed to hold a belief depending upon the quantity and quality of the available evidence.

15 As Susan Haack notes:

What is distinctive about natural-scientific inquiry isn’t that it uses a particular mode or modes of inference, but the vast range of helps to inquiry scientists have developed, many of them – specific instruments, specific kinds of precaution against experimental error, specific models and metaphors – local to this or that field or sub-discipline.

theories. A substantial part of having what passes for scientific expertise in a field is an ability to understand and use the tools of the trade.

Susan Haack divides these aids to understanding into several categories, including helps to the senses and helps to reasoning. Instruments that expand our senses are at the very heart of progress in physics, astronomy, chemistry, and biology as well as practical disciplines such as medicine and engineering. Aids to reasoning are also critical. These include mathematics in its many different forms as well as experimental and quasi-experimental designs and other investigatory devices designed to assist in making causal assertions.

For purposes of this article, I focus on the second set of aids: aids to reasoning. Most of the common mathematical and logical aids to reasoning employed by science are designed to facilitate not simply inquiry, but inquiry of a certain type: inquiry into general laws or principles. This does not mean that scientists are uninterested in the particular case; many scientists and individuals in fields that rely on science, such as engineering, devote most of their energy to specific situations. But the heroes of science are those who are able to put forth explanations in terms of general laws that explain a myriad of particular observations.

This interest in the general and the generalizable leads to a second component of scientific conventions, the lack of a timetable. An inquiry takes as long as it takes and with respect to many questions the answer experts are most comfortable with is, “we don’t know.” “We don’t know” does not necessarily mean that we don’t have a guess. Often it means we do not

16 Id. at 98.
17 See William R. Shadish, Thomas D. Cook & Donald T. Campbell, Experimental and Quasi-Experimental Designs for Generalized Causal Inference (2002). Even the social sciences play an important role in this process when they uncover and document the many systematic reasoning errors that result from judgment by heuristics and then suggest affirmative steps we might take to minimize such errors. See Cass R. Sunstein, Cognition and Cost-Benefit Analysis, 29 J. LEGAL STUD. 1059 (2000).

A third set of aids discussed by Haack are helps to evidence sharing and intellectual honesty. This includes things such as peer review, publication, replication of findings and other formal and informal devices that involve scientists looking over each other’s shoulders. Often this peek over the shoulder focuses on the correct use of the first two types of aids; instruments, mathematics and experimental design. All of these aids are fallible and none guarantees that we will arrive at correct outcomes. From the supposed benefits of bleeding to the more recent realization that many ulcers have a bacteriological, not a psychological source, it is easy to point to occasions where we have been lead astray for lengthy periods of time. Collectively, however, these conventions are thought to facilitate inquiry over the long term.
have enough evidence of the kind we find persuasive to support a conclusion about a phenomenon.\textsuperscript{18}

If science tends to focus on systematic and general knowledge, most trials must deal with specific events. To be sure, law often is concerned with general questions, e.g., does drug X cause injury Y? However, it is also nearly always concerned with what happened to a specific person at a particular point in time. Did drug X cause the plaintiff's injury? Were these fingerprints left at the crime scene those of the defendant?

This inquiry into individual causation is accompanied by a mindset that is contrary to the “wait and see” attitude of science. Legal conventions ask experts to make a decision now based on the evidence at hand.\textsuperscript{19} Fred Prichard quotes the following passage from an expert confronting for the first time the law’s push for a decision now.

Bill (the expert’s attorney) asked me a question about whether the belt was on or not, the lap belt. And I said, “Well, could have been. But then, it may not have been.” Woo, rockets went off. “What do you mean? You're my expert in this case, and you say it 'could be' or 'couldn't be'?” Look, I'm going to tell you. The other side doesn't waffle. They pick one view. And they will push that view. And they will make their case in front of a jury. And there will be no misunderstanding. There will be no gray area. They will take a position one way or the other and make it stick. Now, they don't have any other course of action. That's their life. They make their living going in front of juries and making statements, whether they have facts to back them up or not. Now you, you can go back to designing cars. You have another career. They don't. You better start thinking like they do.\textsuperscript{20}

This anecdotal evidence is supported by survey research. Champagne et al. report that 56% of the experts they

\textsuperscript{18} When a community of investigators say this they are referring to evidence which is derived from the application of the aids to reasoning (and the instruments) to which a field of inquiry is committed. There may not be a “scientific method” writ large, but there are methods and aids to inquiry to which communities of scholars are committed and evidence derived from these techniques enjoy greater warrant in the community than other types of evidence. Over time, these methods may change as new ways to collect and observe are created. Methods are, ultimately, simply tools and a new problem may call for new tools. In this sense, methodology is a pragmatic search for what works. However, at any given point in time the ability of an investigator to persuade her peers about some hypothesis without the use of these devices and methods is limited.


interviewed say their lawyers ask them to be less tentative.\textsuperscript{21} They found a similar percentage (57\%) in a second, follow-up study.\textsuperscript{22} Perhaps more alarmingly, 12\% of the experts in the first study and 22\% in the second study agreed with the statement that lawyers try to get their experts to testify to issues for which there is no scientific basis.\textsuperscript{23}

How then should we summarize the conventions of science and law? Three scientific conventions are particularly relevant to this discussion: a) searching for the general and theoretical, b) doing so by employing the methods and techniques accepted by one’s field, and c) an attitude of agnosticism that encourages waiting for persuasive evidence before making up one’s mind.\textsuperscript{24} On the other hand, legal conventions: a) often focus on the specific event and b) push witnesses toward arriving at a conclusion.\textsuperscript{25}

For the moment, I wish to set aside the second difference, science’s wait-and-see attitude versus the law’s desire to arrive at a conclusion, and focus on the first difference. The fact that much of science focuses on understanding the general, while law is usually interested in the specific, does not always present difficulties. In some areas, the translation from the general to the specific is so well understood that one can reach nearly unanimous consensus about the cause of a specific event through the application of general principles. Engineering is often a case in point. For example, there were multiple hypotheses as to why the I-35 bridge across the Mississippi River at Minneapolis suffered a catastrophic collapse on August 1, 2007.\textsuperscript{26} They included metal

\begin{footnotesize}
\begin{enumerate}
\item Id.; Champagne et al., supra note 21, at 385.
\item A fourth scientific convention is “a commitment to sharing data, intellectual honesty, and disinterestedness.” Joseph Sanders, Science, Law, And The Expert Witness, 72 L. & CONTEMP. PROBS. 63, 66 (2009). I discuss this convention and how it differs from the legal view of expert knowledge as a partisan resource. I do not explore this difference in the present article.
\item Note that the second element in scientific conventions, a commitment to the methods and techniques of inquiry accepted by one’s field, has no direct parallel in legal conventions. I return to this point later in the article. See infra Part VI.
\end{enumerate}
\end{footnotesize}
fatigue among other options. However, in November 2008, the National Transportation Safety Board concluded that the primary source of the failure was a design flaw. The bridge gusset plates were approximately half as thick as they should have been. This failure, combined with substantial increases in the weight of the bridge from earlier modifications and the storage of tons of construction material on the bridge at the time of collapse, caused the failure. Engineers were able to work from the general, e.g., the load bearing capacity of various materials and designs, to the specific flaw in this particular bridge.

Unfortunately, this easy ability to understand the general case and then to translate from the general to the specific case is precisely what is absent in many civil and criminal cases. Whatever the evidence concerning the general principle, e.g., that Vioxx causes heart problems or that individual fingerprints are unique, translating this to the particular case is fraught with difficulty. In the next two sections, I discuss this problem first in the civil context and then in the criminal context.

III. ADMISSIBILITY STANDARDS IN TOXIC TORT CASES

In no area has the Daubert revolution had a greater effect than in toxic torts. The number of cases in which expert causation testimony has been excluded must by now run into the thousands. Many commentators have reacted negatively to this trend, arguing that the bar has been set too high.


27 Id.

28 See NTSB Press Release, supra note 26; see also Matthew L. Wald, Bridge Collapse is Laid to Design Flaw, N.Y. TIMES, Nov. 14, 2008, at A19.


30 Id.

31 This ability is not absent in all cases. DNA testing is an area where the transition from the general to the individual case is well understood. This is also true with respect to toxic torts that produce signature diseases.

32 See supra note 12 and accompanying text. This is especially troublesome to critics when the exclusion of the expert testimony results in a summary judgment for the defendant because the plaintiff no longer has any admissible evidence on causation.

Clearly, the bar is higher than it once was. Before Daubert, very few cases were concluded as a result of an expert witness admissibility determination. The Bendectin litigation, of which Daubert is a part, is a good example. Almost all the twenty-five or so Bendectin cases that were heard on the merits were tried to either a judge or a jury. Although the plaintiffs never prevailed in any of these cases, this was
I do not engage in this debate. Rather, I argue that the height of the bar depends on the causal question being addressed. The causal question in toxic tort cases is usually divided into two parts: general causation and specific causation. The general causation question is whether a substance or drug has been shown to harm any individuals. The specific causation question is whether the harm suffered by the plaintiff was caused by the substance or the drug in question. When evidence is excluded, is it usually because of a failure to present reliable evidence on general causation or a failure to present reliable evidence on specific causation?

Unfortunately, the question is easier to pose than it is to answer. Within the toxic tort arena, most specific causation testimony is presented as “differential diagnosis” testimony. But one cannot judge the frequency with which specific causation testimony is excluded simply by looking at the frequency with which “differential diagnosis” testimony is excluded. This is because most Daubert opinions rule on general causation before reaching the question of specific causation. They require plaintiffs to “rule in” the alleged causal agent, i.e., they must show that the agent causes the injury to some individuals, before “ruling out” other possible causes.

The process of differential diagnosis is undoubtedly important to the question of “specific causation.” If other possible causes of an injury cannot be ruled out, or at least the probability of their contribution to causation minimized, then the “more likely than not” threshold for proving causation may not be met. But, it is also important to recognize that a fundamental assumption underlying this method is that the final, suspected “cause” remaining after this process of elimination must actually be capable of causing the injury. That is, the expert must “rule in” the suspected cause as well as “rule out”
Specific causation witnesses are frequently excluded because neither they nor some other expert has provided sufficient evidence to “rule in” the suspect substance. 36

Once we set those cases aside, that is, once we look only to the cases where there is evidence of general causation, what evidence must the expert present in order to survive an admissibility challenge to the specific causation testimony? The answer is, not very much.

In some Frye jurisdictions, experts conducting a differential diagnosis are considered to be “experience experts” 37 and are allowed to testify without any reliability filter. 38 The

other possible causes. And, of course, expert opinion on this issue of “general causation” must be derived from a scientifically valid methodology.


The great majority of all federal cases on point come to the same conclusion. See 3 FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE § 21.6 (2009).

36 Sometimes, the “ruling in” analysis focuses on the question of dosage. To how much of a substance was the plaintiff exposed? This, too, is best thought of as a question of general causation. See id.


Experience evidence is not restricted to civil cases. As several scholars have noted, much of the testimony of forensic experts may be viewed as experience evidence. See Simon A. Cole, Toward Evidence-Based Evidence: Supporting Forensic Knowledge Claims in the Post-Daubert Era, 43 TULSA L. REV. 263, 276 (2007); Lyn Haber & Ralph Norman Haber, Experiential or Scientific Expertise, 7 LAW, PROBABILITY & RISK 143 (2008). That is, the experts justify their opinion on the basis of their experience with respect to the task at hand.

38 For example, in Kuhn v. Sandoz Pharmaceuticals, the plaintiff offered expert testimony that his wife’s death was “direct[ly] and proximately cause[d]” by her ingestion of Parlodel, a drug taken to suppress lactation in women who chose not to nurse their newborn children. 14 P.3d 1170, 1173-75 (Kan. 2000). The defendant challenged the admissibility of this testimony. Id. at 1180. The Kansas Supreme court held that the Frye test is not applicable to cases where the expert offers “pure opinion” testimony. Id. at 1178. “The validity of pure opinion is tested by cross-examination of the witness.” Id. at 1179. The plaintiff’s three experts offered to testify that Parlodel caused or contributed to Bishop’s death. Id. at 1175. They arrived at this result through a process of “differential diagnosis” by which they considered and ruled out other causes. Id. at 1177. Apparently, Kuhn removes most if not all medical doctor differential diagnosis testimony from any judicial reliability assessment.

Florida also has adopted this position. In Marsh v. Valyou, echoing Kuhn, the court concluded that “[i]t is well-established that Frye is inapplicable to ‘pure opinion’ testimony. . . . [b]ecause testimony causally linking trauma to fibromyalgia is based on the experts’ experience and training, it is ‘pure opinion’ admissible without having to satisfy Frye.” 977 So. 2d 543, 548-49 (Fla. 2007).

The Arizona Supreme Court adopted a similar approach to expert testimony on repressed memory in Logerquist v. McVey, 1 P.3d 113, 123 (2000). The Logerquist court said:
federal and state courts that have adopted Daubert, refuse to adopt this approach. However, as the following discussion indicates, they do not apply a strenuous admissibility standard for specific causation testimony.

What are the admissibility standards? First the expert must offer more than simple temporal order, i.e., the injury followed the exposure and, therefore, the exposure caused the injury. Nevertheless, experts continue to offer this as proof of causation. The very fact that temporal order is so frequently the basis of exclusion is itself an indication of the relatively low threshold set for the admissibility of differential diagnosis testimony for in the context of many long latency period toxic torts temporal order is nearly no evidence at all. Plaintiff experts may also be excluded if they fail to address and rule out

Although compliance with Frye is necessary when the scientist reaches a conclusion by applying a scientific theory or process based on the work or discovery of others, under [Arizona Rules of Evidence 702 and 703] experts may testify concerning their own experimentation and observation and opinions based on their own work without first showing general acceptance.


A few Daubert opinions have come close. For example, in Emig v. Electrolux Home Products, Inc., the court said:

“It is not appropriate to invoke the Daubert test in cases where expert testimony is based solely on experience or training, as opposed to a methodology or technique.” Indeed, where the expert’s opinion is based on “years of accumulated learning and insight,” the reliability of such opinion “should be assessed without resort to the Daubert factors.”


See 3 FAIGMAN, ET AL., supra note 35, at § 21:7. See, e.g., Whiting v. Boston Edison, Co., 891 F. Supp. 12, 23 (D. Mass. 1995) (“[Plaintiff’s experts] propound the argument that . . . because Gary Whiting was exposed to radiation before he contracted [acute lymphocytic leukemia], his ALL must have been caused by radiation exposure. This is a classic illustration of the logical fallacy post hoc ergo propter hoc.”); Schmaltz v. Norfolk & Western Ry. Co., 878 F. Supp. 1119, 1122 (N.D. Ill. 1995) (“It is well settled that a causation opinion based solely on a temporal relationship is not derived from the scientific method and is therefore insufficient to satisfy the requirements of Fed. R. Evid. 702.”).

There are a handful of cases that balk at even this limitation. See Kannankeril v. Terminis International, Inc., 128 F.3d 802 (3d Cir. 1997).

Even some of the temporal order cases are, at bottom, about general causation. See Ervin v. Johnson & Johnson, Inc., 492 F.3d 901 (7th Cir. 2007).
other possible causes. Experts who fail to consider alternatives are often excluded.  

If, however, experts do avoid these two obvious inadequacies, their testimony is rarely excluded. When doctors employ standard diagnostic techniques many courts are likely to admit their differential diagnosis testimony. The classic statement of this position comes from the In re Paoli opinion: “to the extent that a doctor utilizes standard diagnostic techniques in gathering . . . this information, the more likely we are to find that the doctor’s methodology is reliable.” Moreover, most courts would agree with Paoli that a failure to account for all possible causes does not render expert opinion based on differential diagnosis inadmissible.  

John’s Heating Service v. Lamb, is a state court opinion supporting this position:

Of course, “[a] differential diagnosis that fails to take serious account of other potential causes may be so lacking that it cannot provide a reliable basis for an opinion on causation.” But here that is not the case. Here, doctors experienced with carbon monoxide exposure performed the differential diagnosis, which included making physical examinations, taking medical history, and reviewing clinical tests. In addition, the diagnosis was bolstered by a temporal relationship between the symptoms and the possible carbon monoxide exposure and the discrepancy between Cynthia’s performance and verbal IQs corresponding almost uniquely to carbon monoxide poisoning. An expert’s causation conclusion should not be excluded because she has not ruled out every possible alternative; rather, existing possible alternatives should affect the weight that the jury gives the experts’ testimony.

---

42 See Terry v. Ottawa Cty. Bd. of Mental Retardation & Developmental Delay, 658, 847 N.E.2d 1246 (Oh. Ct. App. 2006) (“We agree with the trial court: Dr. Bernstein did not conduct a scientifically valid differential diagnosis, because his method relied primarily upon temporal relationships and because he did not rule out other possible causes. He was properly barred from testifying to specific causation.”); see also Roche v. Lincoln Property Co., 278 F.Supp.2d 744 (E.D. Va. 2003) (Plaintiff’s expert’s testimony that mold caused the plaintiffs’ allergy-like symptoms excluded in part because he failed “to rule out the Roches’ significant allergies to cats, dust mites, grasses, weeds, and trees as potential causes for the Roches’ symptoms.” Mrs. Roche had been to the emergency room on several occasions with similar symptoms prior to moving to the defendant’s apartment.).

43 In re Paoli R.R. Yard PCB Litigation, 35 F.3d 717, 758 (3d Cir. 1994).


The process of ruling out other causes is relatively easy when there is a unique connection between an exposure and a disease (a so-called signature disease)\textsuperscript{47} or, as in \textit{John's Heating}, when there is a unique connection between an exposure and a particular set of symptoms. The task becomes more difficult when various causes do not produce demonstrable differences in the disease.

The task becomes even more difficult when the causes of the type of injury under investigation are not well understood. When the etiology of an illness is not well understood, there will be many idiopathic injuries. This raises an important general point. Should courts permit experts to present differential diagnosis testimony when the clear weight of scientific evidence points to the fact that the substantial majority of certain types of injuries are from unknown causes? Logically, when this situation arises the best differential diagnosis would be legally insufficient. For example, if, with respect to some injury, we know that 5\% of the cases are caused by an exposure to a drug, 5\% are caused by another known cause, and 90\% have no known cause, then even if a differential diagnosis clearly excludes the other known cause it remains the case that it is much more likely than not that the cause in any individual case is not the drug.

\textit{Doe v. Ortho-Clinical Diagnostics, Inc.},\textsuperscript{48} apparently adopted this line of argument. \textit{Doe} is one of a number of cases involving the question of whether exposure to the mercury in thimerosal, a preservative once used in vaccines and other biologic products, is capable of causing autism in children.\textsuperscript{49} After excluding the plaintiff's causal expert on general causation grounds, the court made the following observation:

More troubling, however, is that Dr. Geier's differential diagnosis failed to acknowledge the one conclusion that is generally accepted in the medical community with respect to the causation of autism, which is, that its cause is genetic, but that the exact genetic sequence of autism is unknown. . . . Although Dr. Geier apparently

\textsuperscript{47} This is the situation with respect to asbestos exposure and mesothelioma.  
has considered a number of specific genetic disorders in performing his differential diagnosis, the Court finds that his failure to take into account the existence of such a strong likelihood of a currently unknown genetic cause of autism serves to negate Dr. Geier’s use of the differential diagnosis technique as being proper in this instance.\textsuperscript{50}

What is interesting about \textit{Doe} is that its position is quite rare. Courts do not generally rule against the admissibility of a differential diagnosis on these grounds and note that even in \textit{Doe} the passage was a bit of an afterthought, perhaps even dicta. Recall, the court had already concluded that the plaintiff could not prevail on general causation.

In sum, the high exclusion rate in toxic torts disguises the fact that these cases address two separate issues, general and specific causation. With respect to general causation, the courts rely heavily on the available science to exclude expert testimony. However, with respect to specific causation, courts are, on balance, much more lenient. If an expert follows normal procedures, i.e., collects a medical history of the patient and the patient’s family, conducts appropriate laboratory and diagnostic tests, and gives consideration to other causes, and if there is no a glaring alternative cause that seems far more likely to have caused the illness,\textsuperscript{51} courts are likely to admit differential diagnosis testimony even though there simply are no tests one can perform that will produce a quantified estimate as to whether an injury is the result of one cause or another.\textsuperscript{52}

IV. ADMISSIBILITY STANDARDS IN FORENSIC CASES

Prior to \textit{Daubert}, the \textit{Frye} “general acceptance” test was used primarily in criminal cases.\textsuperscript{53} Courts excluded proffered

\textsuperscript{50} \textit{Ortho-Clinical}, 440 F. Supp. 2d at 477 (internal citations omitted).


\textsuperscript{52} \textit{Susan R. Foulter, Science and Toxic Torts: Is There a Rational Solution to the Problem of Causation?}, 7 HIGH TECH. L.J. 189, 209-10 (1992) (“Scientists know very little about how, in a mechanistic sense, toxic substances cause disease such as cancer or birth defects.”).

\textsuperscript{53} Indeed, nearly all \textit{Frye} admissibility rulings that excluded expert testimony were in the criminal area. Federal civil cases in which the \textit{Frye} test was employed to exclude testimony are extraordinarily rare. Kenneth Chesebro, \textit{Galileo’s Retort}, 42 AM. U. L. REV. 1637, 1695 (1993) reports there were only two federal appellate court opinions excluding civil evidence on \textit{Frye} grounds prior to the lower court \textit{Daubert} decision. Christophersen v. Allied-Signal Corp., 939 F.2d 1106, 1115-16 (5th Cir. 1991) (employing \textit{Frye} to determine that district court was within its discretion to exclude medical expert’s testimony in civil case where that testimony was not generally accepted within relevant scientific community), cert. denied, 112 S. Ct. 1280 (1991) (using \textit{Frye} to exclude epidemiological re-analysis studies in civil suit); \textit{Barrel of Fun, Inc. v. State Farm Fire & Casualty Co.}, 739 F.2d 1028, 1031 (5th Cir.
expert testimony in a few areas. The inadmissibility of lie
detector results is the most well known example.54 With respect
to most areas of forensic evidence, however, the courts
concluded that expert testimony met the general acceptance
standard.55 As pointed out by Paul Giannelli, the courts arrived
at this result by defining the relevant field of expertise as the
very group of individuals, e.g., handwriting analysts, whose
expertise was being judged.56 Unsurprisingly, this group
vouched for its own endeavor.7 Admissibility rulings did not
concern themselves with the quality of the evidence.

Daubert offered an opportunity to revisit these
admissibility decisions but by and large the pattern of liberal
admissibility decisions has not changed. Although federal
courts and state courts in jurisdictions that have adopted
Daubert often pay lip-service to that test, it is rarely used to
exclude forensic evidence experts.58

There are, undoubtedly, many reasons why courts apply
liberal standards when considering forensic evidence.59
However, I believe an important factor is the same one that
leads to liberal admissibility rulings in toxic tort, specific
causation determinations. There is very little hard scientific
evidence upon which to base opinions. As a consequence,
similar to the specific causation situation, the courts rely on

---

54 See 4 FAIGMAN ET AL., supra note 35, at § 35.
55 See 1 FAIGMAN ET AL., supra note 35, at § 1:35.
57 Id. at 1203.
58 For examples of this reluctance to apply Daubert factors, see 4 FAIGMAN ET
AL., supra note 35, at § 32 (fingerprint identification), § 33 (handwriting identification),
§ 36 (bitemark identification).
59 They include: a) bias in favor of the state in criminal prosecutions, b) the
quality of criminal defense attorney and their limited resources, c) the negative effects
on judges if they refuse to permit the state to present its experts (a consideration that
may be particularly relevant to elected judges, d) the long history of admissibility of
forensic evidence which creates its own inertia to keep things the way they have
always been, and e) judicial belief of the claims of forensic experts. For a very useful
discussion of why lower courts may actually believe forensic experts despite the limited
empirical evidence supporting some of their claims, see Michael J. Saks, Explaining the
Tension Between the Supreme Court's Embrace of Validity as the Touchstone of
Admissibility of Expert Testimony and Lower Courts' (Seeming) Rejection of Same, 3
expert testimony that is not much more than a clinical judgment from an experience-based expert.  

Before I expand on this point, I should be clear on another. I do not mean to suggest that the causal issue in the forensic context is exactly the same as the issue in toxic torts. Within the context of forensic experts, the distinction between general and specific causation is rarely, if ever made. We simply do not think about forensic evidence in that way. Indeed, the terms “general causation” and “specific causation” seem to be terms of art that are restricted to the toxic tort context.

Nevertheless, one can think of forensic evidence in these terms. From one point of view, the connection between general causation and specific causation is easier in forensic cases. The problem in toxic torts is to determine which of several things caused the plaintiff’s disease. This problem does not exist in forensics. Latent fingerprints are “caused” by fingers, not some other source. In the jargon of toxic torts, latent prints are a “signature disease” caused by human fingers. This is true of even the most suspect forms of forensic evidence, e.g., bite marks. It is, in fact, the existence of this relationship that gives much forensic evidence testimony its initial plausibility.

Another, perhaps more useful way to compare the causal question in toxic tort and forensic cases is to think of each individual as a source of many latent prints. That person is, in the jargon of toxic torts, an entity that is a “general cause” of some prints. The individual problem is determining whether this person, rather than all the other plausible sources of the print, is the “cause” of a particular print.  

From this point of view, the forensic evidence problem is similar to the problem that arises in tort when we have an indeterminate defendant. This issue was brought into starkest relief in the DES cases. See Sindell v. Abbott Laboratories, 26 Cal.3d 588, 601 (1980). In those cases, through no fault of their own, the plaintiffs could not identify which drug company manufactured the drug sold to their mothers many years earlier. In the Sindell case the court solved the plaintiff’s proof problem by forcing defendants to prove that they did not sell the prescription or be liable for a share of the plaintiff’s injury based on each defendant’s market share of sales of DES in the relevant market. This solution has rarely escaped the narrow confines of the DES cases, and generally a plaintiff simply cannot prevail in this circumstance. Id.

A related problem arises when many defendants contributed some part of the total dose of a harmful substance that caused the plaintiff’s injury. This issue arises in asbestos cases. One solution is to ask the jury to assign liability to each defendant based on the percentage of the total risk of injury generated by that defendant’s product. Thus, if a defendant made 10 percent of the total asbestos to
In both areas, the accuracy with which we can make this judgment turns on the quality and quantity of our general understanding of a phenomenon and our ability to translate this knowledge to a judgment about the individual case. In the forensics arena, as in the toxic tort arena, the courts must try to answer a question without the benefit of much underlying general scientific knowledge. The problem in toxic torts is that the base rate information supplied by epidemiological research and, to a lesser extent, animal studies, is not easily translated into information about causation in the particular case. The problem in most forensics situations is in one sense more fundamental: with the exception of DNA testing, there is very little base rate information at all. In the next few paragraphs, I illustrate the problem using the example of fingerprint identification.

A. Fingerprint Identification

The primary method used to examine fingerprints in the United States is called ACE-V, an acronym which stands for the stages of the examination: Analysis-Comparison-Evaluation-Verification. As described by Haber and Haber, the ACE-V examination proceeds as follows.

At the Analysis Stage, “the fingerprint examiner looks at a latent fingerprint and decides whether it contains sufficient quantity and quality of detail so that it exceeds the which the plaintiff was exposed, it would ideally be assigned 10 percent of the liability. See Rutherford v. Owens-Illinois, Inc., 16 Cal. 4th 953, 957 (1997).

It is important to note that in both the Sindell and the Rutherford situations the courts are prepared to adopt this solution only in the cases where every one of the defendants has breached a duty to the plaintiff. Obviously, this is never the situation in criminal cases. One potential defendant “breached a duty” to the plaintiff, but all the other potential defendants are innocent.

On rare occasions, one might have a specific causation issue in forensics which is similar to that in toxic torts. For example, a case might arise where the question was whether a bite mark was caused by human or animal teeth.

I chose fingerprint identification because it is an area of forensics with high face validity. If a problem exists here, a fortiori, it exists in other areas such as handwriting analysis.

Lyn Haber & Ralph Norman Haber, Scientific Validation of Fingerprint Evidence Under Daubert, 7 LAW, PROBABILITY & RISK 87, 87 (2008).

Id.
standard for value.” If it does not, no further steps are possible and the prints are rejected.

If a latent print does meet the value standard, the examiner collects as much evidence as possible on “the nature of the surface on which [the print] was deposited, the amount and direction of pressure used in touch[ing the surface, and the way] . . . in which the ridge details of the [print] were transferred onto the surface,” e.g., sweat. All of this is employed in the next step to account for inevitable distortions between the latent print and the print against which it is being compared.

For prints that do meet the value standard, the examiner chooses a feature-rich area of the latent print. “Within this area, he selects the particular features along the various ridge paths in the latent print, in their spatial locations relative to one another . . . .”

In the Comparison Stage, the examiner attempts to ascertain whether one of the suspects’ fingers made the latent print by comparing the same area that he chose for the latent print. Failure to find a correspondence that cannot be accounted for by factors such as distortion lead to an exclusion of that finger. This, of course, is the most likely result. If there are sufficient points of similarity and no excluding differences, the examiner will return to the latent print, examine another area of the latent print, and then again compare this area to the exemplar print. This process will be repeated until “all the features of the latent print have been compared.” If there are sufficient points of comparison and no excluding differences, the examiner proceeds to the Evaluation Stage.

“In [the Evaluation Stage], the examiner applies a sufficiency standard to the amount of corresponding agreement between the latent and the exemplar [print].” “If the amount of corresponding agreement exceeds the sufficiency standard,
then the examiner concludes that the crime scene latent print [was made by the suspect finger].”77 (In the language of latent print examiners, “the crime scene latent print can be individuated to the suspect.”)78 “If the amount of agreement does not reach th[is] standard,” the examiner may say the comparison is inconclusive.79 This sufficiency standard can be numeric, i.e., there are X number of points of similarity between the latent print and the exemplar.80 Or, the standard can be experiential, “based on the individual examiner’s training and experience.”81 Fingerprint examiners in the United States rarely use a numeric standard.82

In the Verification Stage, which is done in larger laboratories, after an examiner has concluded that the latent print did come from a suspect finger, a second examiner confirms or disconfirms the conclusion.83 The verification is not a complete re-analysis, but rather is a review of the evidence and the conclusions of the examiner.84

As cautious and thorough as this sounds, apparently there is no peer-reviewed study testing the validity of the ACE-V method.85 Indeed, such a study would be difficult to conduct because the details of how an examiner is to proceed are left quite open ended.86 As a consequence, here is what a recent National Academy of Science (the “National Academy”) report on the state of forensic evidence has to say about ACE-V.

[T]he assessment of latent prints from crime scenes is based largely on human interpretation. Note that the ACE-V method does not specify particular measurements or a standard test protocol, and examiners must make subjective assessments throughout. In the United States, the threshold for making a source identification is deliberately kept subjective, so that the examiner can take into account both the quantity and quality of comparable details. As a

77 Id.
78 Id. at 91.
79 Id.
80 Id.
81 Haber & Haber, supra note 63, at 91.
82 Id. at 102.
83 Id. at 91.
84 Id.
85 Id. at 95.
86 NAT’L RESEARCH COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 139 (2009).
result, the outcome of a friction ridge analysis is not necessarily repeatable from examiner to examiner . . . .

This subjectivity is intrinsic to friction ridge analysis, as can be seen when comparing it with DNA analysis. For the latter, 13 specific segments of DNA (generally) are compared for each of two DNA samples. Each of these segments consists of ordered sequences of the base pairs, called A, G, C, and T. Studies have been conducted to determine the range of variation in the sequence of base pairs at each of the 13 loci and also to determine how much variation exists in different populations. From these data, scientists can calculate the probability that two DNA samples from different people will have the same permutations at each of the 13 loci.

By contrast, before examining two fingerprints, one cannot say a priori which features should be compared. Features are selected during the comparison phase of ACE-V, when a fingerprint examiner identifies which features are common to the two impressions and are clear enough to be evaluated. Because a feature that was helpful during a previous comparison might not exist on these prints or might not have been captured in the latent impression, the process does not allow one to stipulate specific measurements in advance, as is done for a DNA analysis. Moreover, a small stretching of distance between two fingerprint features, or a twisting of angles, can result from either a difference between the fingers that left the prints or from distortions from the impression process. For these reasons, population statistics for fingerprints have not been developed and friction ridge analysis relies on subjective judgments by the examiner. Little research has been directed toward developing population statistics, although more would be feasible.

As the National Academy report notes, population statistics are a perquisite to fully quantifying the diagnosticity of an observed set of fingerprint characteristics. If the characteristics are rare, then a reported match is more diagnostic of the claim that two prints came from a common source. For example, in DNA testing, the random match probability, i.e., the frequency with which a genetic profile exists in the population, captures the diagnosticity of a match.

Developing population statistics for fingerprints will not be an easy task. We do not have a “map” of fingerprints in the same sense that we have a map of the human genome. Thus,

---

87 Id. at 139-40. The report cites recent research that suggests even “experienced examiners do not necessarily agree with even their own past conclusions when the task is presented in a different context” at a later time. See id. (citing Itiel E. Dror & David Charlton, Why Experts Make Errors, 56 J. FORENSIC IDENTIFICATION 600 (2006)).

88 Id.

we do not know the frequency of different patterns in the way that we know the variations in the sequencing of base pairs in specific loci. Nor do we have strong evidence of the independence of patterns.\textsuperscript{90}

Even if we were to have such base rate information, we would find it to be of less practical use than it is in DNA testing. In most cases, DNA matches are made with sufficient DNA that we can make comparisons at multiple locations.\textsuperscript{91} The DNA samples contain complete or nearly complete information about the sequences of base pairs at multiple loci. Metaphorically, the samples are nearly “perfect prints.” The problem in the real world of fingerprint identification arises when the latent print is far, far from perfect. The impression left by a given finger will differ every time because of variations in pressure and the impression medium.

Given this reality, the National Academy report proposes some other steps that could be taken to move the area toward a sounder scientific footing.\textsuperscript{92} For example, the field could conduct research on the variables that effect latent print differences.\textsuperscript{93} Note that these steps are designed to improve our general understanding of fingerprint marks, a general causation-like issue. Even more helpful in the short run would

\textsuperscript{90} Independence means that the sequence of base pairs at one location is independent of the sequence at another location. When independence exists, one can use the product rule to determine the overall probability of a match. Thus if a particular sequence at location one occurs in 5% of the population and a particular sequence at location two occurs in 10% of the population, the probability of observing the same sequence at both loci is 5% times 10% or 0.5%. In the early years of DNA testing it was not clear that the independence assumption was valid, causing some to propose a more conservative rule when combining probabilities. See Richard Lempert, \textit{DNA, Science and the Law: Two Cheers for the Ceiling Principle}, 34 JURIMETRICS J. 41, 42-43 (1993).

\textsuperscript{91} The widespread adoption of Polymerase Chain Reaction (PCR) test protocols has greatly reduced the occasion where the sample quantity of DNA is insufficient for a proper analysis. See 4 PAIGMAN ET AL., supra note 35, at § 30:44.


\textsuperscript{93} NAT'L RESEARCH COUNCIL, supra note 86, at 105-06.
be research on reliability. The most frequently suggested way to achieve better reliability estimates is to engage in serious proficiency testing. Such testing would provide a better estimate of the error rate (the frequency of false positives and false negatives) associated with reported matches.

Some fingerprint identification proficiency tests have occurred, but they have not used rigorous double blind methodology and most observers agree that the tests have not involved challenging partial prints. Nearly everyone agrees that reliability is very high when the examination involves good-quality impressions of all ten fingers. The critical question is how far we can move from this best-case situation before identification accuracy begins to deteriorate. Proficiency tests employing challenging partials address this question. Without some studies along these lines we are unable to assess the reliability of expert assertions in different identification settings.

Most courts have not been sensitive to the importance of distinguishing among different identification situations. With very, very few exceptions, they have admitted fingerprint identifications. Courts have refused to conduct Daubert hearings, have implicitly reversed the burden of persuasion to require the defendant to demonstrate that a fingerprint identification...
identification is not reliable, have refused to focus on the “task at hand” as required by *Kumho Tire*, have refused to conduct an assessment of the evidence on fingerprint reliability, have admitted expert testimony by relying on the fact that other courts have admitted the testimony, have relegated any concerns about validity to weight, not admissibility, and in general have lowered the bar to the level necessary to admit fingerprint identification.

The story is the same with respect to many other types of forensic evidence. Courts often seem to simply mouth the *Daubert* criteria without actually assessing whether the proffered evidence meets these criteria. In sum, the critique of judicial leniency toward forensic experts is well founded. But is it any more lenient than proof of specific causation in toxic tort cases? The next section compares the two areas.

---

106 Id. A rare exception to this landslide of opinions admitting fingerprint evidence was *United States v. Llera Plaza*, 179 F. Supp. 2d 492 (E.D. Pa. 2002). However, Judge Pollak later withdrew his initial opinion excluding fingerprint evidence and admitted the evidence. *United States v. Llera Plaza*, 188 F. Supp. 2d 549 (E.D. Pa. 2002); *see also* United States v. Crisp, 324 F.3d 261, 272 (4th Cir. 2003) (Michael, J., dissenting).
107 Faigman, supra note 8, at 718; *see also* Johnson v. Commonwealth, 12 S.W.3d 258, 260-64 (Ky. 1999). For a discussion of Johnson, see Michael J. Saks, *Protecting Factfinders From Being Overly Misled, While Still Admitting Weakly Supported Forensic Science into Evidence*, 43 Tulsa L. Rev. 609 (2007):

In a challenge to the admissibility of microscopic hair identification evidence, the Kentucky Supreme Court purported to be conducting an analysis under that state’s version of *Daubert*. The record was devoid of research studies on the validity of asserted microscopic hair identification expertise, so the Court relied entirely on the “general acceptance” criterion of *Daubert*. But there was no evidence of that in the record either. So the Court turned to its own earlier (pre-*Daubert*) Kentucky decisions in search of general acceptance of microscopic hair comparison. But not one of the earlier cases admitting testimony on hair identification said a word about general acceptance of the technique. So the Court stated that it would assume that those earlier decisions must have: addressed the question, conducted an appropriate inquiry, and found general acceptance. How else could they have admitted the testimony?

*Id.* at 619-20. In point of fact, the *Frye* test had not been adopted or even mentioned in Kentucky cases until after those cases were decided. As Saks notes, “this Court had to create out of thin air the basis for admission under the weakest of the *Daubert* prongs.” *Id.* at 620.
V. COMPARING TOXIC TORT AND FORENSIC ADMISSIBILITY STANDARDS

I do not wish to argue that there are no differences in admissibility standards in toxic tort cases and in admissibility standards in forensic cases. There are. Many tort plaintiffs lose their lawsuit at the summary judgment stage after the court has excluded their causation experts. Only very rarely does the same fate befall the state in criminal prosecutions. However, the differences disguise an underlying similarity that is revealed once we control for the quantity and quality of the scientific evidence available to address various causal questions. Plaintiff experts are excluded in toxic tort cases largely because of an inability to show general causation, something about which there is often a body of empirical evidence. However, when there is relatively little evidence, which occurs most frequently with respect to proof of specific causation, the courts set a threshold that is not much higher than that required in the forensic area to prove a particular defendant left some piece of trace evidence at a crime scene.

This does not mean that experts who testify about individual causation always have little or no support for the position they espouse. Quite frequently, they do. The point is well made by Professor Mnookin with respect to fingerprint evidence.

[T]he reality is that there are plenty of cases where the totality of the evidence—including but not limited to a fingerprint identification—leaves little practical doubt about ground truth. Fingerprint identification’s 100 years of use in a variety of contexts does not come close to answering all the questions about precisely how accurate it is, or how commonly identification errors are made, but it does provide some degree of prima facie evidence of its general validity.

The same may be said for many differential diagnoses. Often the probabilistic evidence that a particular substance caused a particular illness may be very high and we can state with substantial certainty that one caused the other even in the absence of a mechanistic understanding of how this occurs.


109 See Mnookin, supra note 96, at 134.

110 I certainly do not mean to suggest that such probabilistic evidence is never a sufficient substitute for a mechanistic understanding. This argument is similar to the cigarette industry’s long standing argument that the causal link between tobacco usage
The point is not that all such testimony is unreliable (although in certain areas such as bite marks, a large percentage of the testimony may be of very low validity). Rather, the point is that the threshold for admissibility is low and, in part, this is a judicial response to the fact that there is little useful quantified empirical evidence courts may employ to establish a higher admissibility baseline.\footnote{In the absence of direct empirical evidence, following a set routine and expressing conclusions in a formulistic way takes on increased importance. In the toxic tort arena, the best differential diagnoses involve factors such as: a patient examination and patient history, diagnostic tests, laboratory tests, tissue samples and biopsies, and genetic testing. See 3 FAIGMAN, ET AL., supra note 35, § 21:31-39. Failure to structure one’s testimony along these lines or a failure verbally to give full consideration to all of the data may lead to the exclusion of the expert’s testimony. See, e.g., Cooper v. Smith & Nephew, Inc., 259 F.3d 194, 203 (4th Cir. 2001); Viterbo v. Dow Chem. Co. 826 F.2d 420, 424 (5th Cir. 1987).
Likewise, fingerprint expertise may be rejected if the state makes no effort to explain and justify an examiner’s conclusion. See Jacobs v. Gov’t of the Virgin Islands, 53 Fed. Appx. 651 (3d Cir. 2002). The recent emergence of the ACE-V terminology for fingerprint analysis may be understood from this perspective. Professor Cole notes that the terminology was first adopted after the Daubert opinion. “Given the fortuitous timing, one might suspect that the term was adopted in the wake of Daubert to lend forensic fingerprint identification a scientific patina.” Simon A. Cole, Grandfathering Evidence: Fingerprint Admissibility Rulings From Jennings to Llera-Plaza and Back Again, 41 AM. CRIM. L. REV. 1189, 1236 n.201 (2004).
Wrapping testimony in the rhetoric of a methodology does not, of course, make it more reliable. As Professor Cole notes with respect to ACE-V, “[I]t indicates little more than looking at two objects and determining, subjectively, whether they originate from a common source, and then repeating this process.” Id.
Professor Mnookin argues that courts should not settle for explanations of method standing alone.
Indeed, the history of the identification sciences in court shows a repeated and dangerous pattern: when judges are provided with an intuitively appealing description of a science of ‘matching,’ they frequently let poorly specified explanations of method substitute for a more searching assessment of validity and reliability . . . .
Judges, therefore, would be well advised to focus on the degree of testing associated with the claims made by experts rather than emphasizing whether the expert has offered a seemingly plausible explanation of her technique and her conclusion.
VI. CONTEXTUALISM AND THE LAW’S APPROACH TO INDIVIDUAL CAUSATION

Many critics of the lax admissibility standards in forensic cases call upon the courts to tighten things up. At the same time, some call upon courts to ease admissibility standards in civil cases. A first reaction to these two requests might be that the courts should search for one appropriate standard that applies to all admissibility questions. A single standard might be epistemologically more satisfying, but given the social objectives of the law, in my opinion, it would be ill advised. A brief discussion of the standard epistemological approach to knowledge will help us to see why this is the case.

The standard approach to the question of “when . . . it is proper to say someone knows something . . . involves the interplay of three primary variables: belief, truth, and justification.” Belief is a person’s subjective position concerning the truth of a proposition. Truth is the reality of the proposition independent of belief. Justification involves the quality of the reasons for a belief. To count as knowledge, something must be believed to be true, it must be true, and a person’s belief that it is true must be justified. In the absence of belief, we have ignorance. In the absence of truth, we have error. In the absence of appropriate justification, we have mere opinion. What is most noteworthy about this standard approach is that its main concern is not knowledge per se but justification. Even correct beliefs without justification are not knowledge.

For courts, the relevant question is what level of justification should be required before experts are permitted to testify? Clearly, the level must be sufficient to qualify the person as an expert, i.e., in the jargon of evidence law, a person must be in possession of knowledge (and justification for that knowledge) that is beyond the ken of the lay person. But beyond that, how much? Presumably, a single admissibility standard would require the same level of justification from all experts. But from where would such a standard come?

Earlier in the article, I compared legal and scientific conventions. I noted that three relevant scientific conventions

are: a) a search for general and theoretical propositions, b) doing so by employing the methods and techniques accepted by one's field, and c) an attitude of agnosticism that encourages a wait-and-see attitude. I compared these to the relevant legal conventions which: a) often focus on the specific event, and b) push witnesses toward arriving at a conclusion.

The point I wish to make here is that the second element in scientific conventions, a commitment to the methods and techniques of inquiry accepted by one's field, has no direct parallel in legal conventions. Of course, the law has conventional techniques it uses to arrive at causal conclusions. Substantial parts of the law of evidence could be understood in this way. But ultimately when it comes to methods and techniques of acquiring knowledge in the first instance, the law has no methodology of its own. It simply borrows the relevant scientific methodology. Presumably, the law could be indifferent to the methods and techniques used by an investigator or even ask the investigator to forego these techniques when appearing in court. However, most post-Daubert courts are respectful of these methods. Many successful Daubert challenges turn on the argument that the expert failed to use the methods deemed to be appropriate by those in his field when arriving at a causal conclusion.\(^{114}\)

Even if the courts were inclined to establish a single level of justification for causal assertions, what should it be? In some situations, raising the bar may systematically guarantee victory for one side, a result many courts find to be undesirable as a general principle.\(^{115}\) If raising the bar has problems, how about lowering the bar? This does not lead inevitably to a victory for one side, but it has its own problems. If experts are not to use their normal methods for determining causation, what are they to use? Because the question has no obvious

---

\(^{114}\) One need look no further than *Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137 (1999). Referring to the plaintiff's expert, the court said, "Indeed, no one has argued that Carlson himself, were he still working for Michelin, would have concluded in a report to his employer that a similar tire was similarly defective on grounds identical to those upon which he rested his conclusion here." *Id.* at 157.

\(^{115}\) In the toxic tort area, opinions that require the plaintiff to have epidemiological evidence in order to prove causation note that this is not always necessary. *See Merrell Dow Pharm., Inc. v. Havner*, 953 S.W.2d 706, 718 (1997). One suspects that a similar concern underlies many lenient forensic opinions. The court may fear that raising the bar too high would make it impossible for the state to prove its case in many criminal trials.
answer, courts are also reluctant to adopt this position. Where does this leave us?

The epistemological approaches most congenial to the standard legal position on what constitutes adequate justification are contextual approaches to knowledge. The central idea behind contextualism is that the standards for making knowledge attributions vary depending on the context within which they are made. The rules of evidence are applied in a way that is consistent with the contextualist’s fundamental observation that the level of justification we require for something to count as knowledge and coincidentally for someone to be epistemically responsible in asserting a

---

116 I attended a Court of Federal Claims conference in November of 2008. One of the panels was on compensation under the National Vaccine Act. The central topic of the panel was proof of causation under the Act. A number of attorneys from the petitioners’ bar argued for a relaxed standard of causation that would be different from “scientific” causation. Those in favor of this position argued that a medical theory combined with temporal order should be sufficient to prove causation. If one suffered an injury shortly after being vaccinated and some theory supported a causal connection then one should recover under the Act. In support of this position, the petitioners cited Althen v. Sec’y of Health & Human Servs., which set forth three criteria for recovery under the Act: (1) medical theory causally connecting the vaccination and the injury; (2) a logical sequence of cause and effect showing that the vaccination was the reason for the injury; and (3) a proximate temporal relationship between vaccination and injury.

In response to this argument, a special master in a thimerosal case before the court responded:

Althen requires more than merely a medical theory. Petitioners must offer a biologically plausible medical theory.

What is missing from petitioners’ formulation of the medical theory prong of Althen is the requirement that such a theory be reliable.

Under the Vaccine Act, a special master may determine the reliability of a medical theory by considering the framework established by Daubert. See Terran v. Sec’y, HHS, 195 F.3d 1302, 1316 (Fed. Cir. 1999) (framework established by Daubert for evaluating the reliability of evidence appropriate for use by special masters). Daubert requires that an opinion be supported by something more than subjective belief; it must be grounded “in the methods and procedures of science.”


belief, varies according to the context within which the belief is held and expressed. Within the confines of the present discussion, the most important context is the quantity and quality of the available evidence.\textsuperscript{119}

The contextual nature of legal epistemology is both a strength and a weakness. It is a strength because it balances two important legal goals. Earlier in the article, when listing differences between scientific and legal conventions, I noted that the law needs to arrive at a conclusion. A wait-and-see attitude that advises us to bide our time until we arrive at a greater level of justification for a belief may suffice in science, but not in the courtroom. Competing against the goal of resolving cases is the goal of arriving at the factually correct outcome. As Federal Rule of Evidence 102 states, evidentiary rules shall be construed toward the end of ascertaining the truth.\textsuperscript{120} Acquiescing to a low level of justification when in fact experts can do better is to insulate more incorrect outcomes than are necessary. A contextual approach balances these objectives. If, as Aquinas teaches us, prudence is the first virtue, contextualism is indeed a virtue.\textsuperscript{121} And as a practical matter this approach permits the law to sidestep difficult philosophical questions concerning knowledge and get on with the business of deciding cases.\textsuperscript{122}

However, if the contextual approach’s flexibility is its strength, it is also its potential weakness. Without more, it offers no independent standard by which courts may measure the twin problems of epistemological adequacy: when is the

\textsuperscript{119} In this article, I do not propose to provide a complete review of the many ways in which courts adopt a contextual approach. One example, however, might be in order. In the rather well known handwriting analysis, United States v. Starzecpyzel, 880 F. Supp. 1027 (S.D.N.Y. 1995), Judge McKenna was met with a defense challenge that the state’s expert should be excluded because there was little or no scientific support for the reliability of his alleged expertise. The judge concluded that if the \textit{Daubert} standard were applied, the state’s expert should be excluded. Rather than take this step, the judge concluded that \textit{Daubert} did not apply to “skilled” experts. Here, the key point is that the court invents a special category of expert with a special admissibility standard precisely because it recognizes that the standard to which experts were held to in \textit{Daubert} would lead to the exclusion of the state’s document examiner.

\textsuperscript{120} \textit{Fed. R. Evid.} 102.


\textsuperscript{122} One is reminded of Sir Frederick Pollock’s famous aphorism that “The lawyer cannot afford to adventure himself with philosophers in the logical and metaphysical controversies that beset the idea of cause.” \textit{Sir Frederick Pollock, The Law Of Torts} 36 (11th ed. 1920).
admissibility bar set too high and, my primary concern in this article, when is the bar set too low?123

The best answer to date, and an answer consistent with a contextual approach, has been advanced by Professor Nance124 and Professor Mnookin.125 Courts should require a level of justification that is as good as practicably possible. Assuming such a standard is as good as we can do, what does it practically mean with respect to individual causation evidence in toxic tort cases and many types of forensic evidence? Importantly, does it condemn us to perpetually accepting a low threshold when there is little available science? I believe that if we properly understand the nature of the contextual approach, the answer is no. Allow me to elaborate.

If the contextual approach tells courts to require a level of justification that is as good as practicably possible, how should courts understand the context within which to apply this standard? Should the relevant context only be the case at hand and the evidence available at that moment in time, or should we take a longer view? If we focus solely on the case at hand, that is, if we set the level of necessary justification to reflect the state of knowledge at the moment a particular case is being tried, we may find ourselves permanently settling for relatively little by way of justification.126 If, on the other hand,

---

123 With respect to the former question, at least in those areas where experts employ their expertise outside the courtroom, I believe the best answer is to require experts to use the same intellectual rigor they employ in their day-to-day work. See Joseph Sanders, Expert Witness Ethics, 76 FORDHAM L. REV. 1539 (2007). This means that even in the same general area of litigation, e.g. toxic torts, the courts may require more justification when there is substantially more evidence available. This, apparently, is what courts in fact generally do. See David L. Faigman et al., How Good is Good Enough? Expert Evidence under Daubert and Kumho, 50 CASE W. RES. L. REV. 645 (2000).

124 “The best that is reasonably available should be admitted, at least so far as the reliability requirement of Rule 702 is concerned.” Dale Nance, Reliability and the Admissibility of Experts, 34 SETON HALL. L. REV. 191, 241 (2003).

125 “One appropriate focal point for the judge, it seems to me, ought to be whether the expert evidence on offer is as reliable as it can reasonably be, considering the particular context and circumstances.” Mnookin, supra note 96, at 133.

126 This does not mean courts will settle for anything. For example, under some fact patterns even the relatively lenient standard prevalent in specific causation cases may be insurmountable. Consider, for example, Newton v. Roche Labs., Inc., 243 F. Supp. 2d 672 (W.D. Tex. 2002). In this case, the plaintiffs claimed that exposure to Accutane induced their child’s schizophrenia. Excluding the differential diagnosis testimony of the plaintiffs’ expert, the court noted:

Dr. Rossiter’s conclusion that Accutane induced Candida’s schizophrenia relies solely on the temporal proximity of her illness and her taking of the drug. However, he does not adequately consider that Candida’s uncle and sister were schizophrenic and her mother outwardly exhibited symptoms consistent with
we take a longer view, we concern ourselves with the impact of present admissibility decisions on the body of foreseeable future science and foreseeable future cases. I believe that the better course of action is for courts to adopt this longer view.

The longer view is superior because it allows us to recognize situations where the courts may be able to push lawyers and, to the extent possible, science toward better evidence. Moving the system toward better evidence can be accomplished in two ways: by improving the mix of cases and by improving the underlying science. I discuss each in turn.

Improving the mix of cases. At any point in time, an admissibility ruling produces a shadow effect on future cases. A decision that admits expert testimony puts pressure on the other side to mount a more complete case, by including stronger opposing expertise. An exclusionary decision has the opposite effect: it encourages proponents of such cases to produce better evidence in the future.\footnote{Because this article focuses on two areas where admissibility decisions are lax, I focus my attention on the effect of tightening admissibility standards.}

It may be that the nature of the science available on a legal question and the quality of the expert opinions available to report on that science are a constant. In this situation, a higher standard can affect the mix of cases that are litigated, but it cannot influence the quality of future cases. On the other hand, if lawyers have not presented the best possible evidence or have employed less well qualified experts, then a heightened admissibility standard may improve the quality of evidence in the same type of cases.\footnote{There is some evidence that this may have happened in some civil cases.}

Improving the science. In most situations, scientific investigations are, at best, marginally influenced by what is going on in court. When an area of litigation becomes a hot topic, this may spur some research. For example, there is evidence that the litigation surrounding the drug Bendectin generated some research that would not otherwise have been

\footnotetext{27}{Because this article focuses on two areas where admissibility decisions are lax, I focus my attention on the effect of tightening admissibility standards.}

\footnotetext{28}{There is some evidence that this may have happened in some civil cases. Lloyd Dixon & Brian Gill, Changes in the Standards for Admitting Expert Evidence in Federal Civil Cases Since Daubert, 8 PSYCHOL. PUB. POLY & L. 251, 296-99 (2002).}
published.129 Note, however, that it was the litigation itself, not admissibility rulings per se that led to a better scientific understanding of the issue at hand. The influence of legal opinions on the production of science may be much stronger in situations where the legal system itself is the primary market for the information.130

All of this suggests that from the perspective of a longer view, the specific causation in toxic tort cases and in forensic cases is not the same. The impact of admissibility decisions is likely to be greater in forensic cases.

In both areas, admissibility rulings may influence the mix of cases. And marginally, they may influence the quality of expertise presented in a particular trial.131 However, adverse admissibility rulings are likely to have a larger effect on case mix in the forensic setting. This is due to the social organization of criminal prosecutions when compared to the social organization of tort litigations. The latter is much less well organized. Many tort claims are brought by small time entrepreneurial tort lawyers. And even when the plaintiff’s bar is more organized, as it often is with respect to mass torts, there is a much more limited hierarchical structure. Individual lawyers may bring relatively weak cases, unimpeded by concerns for the larger body of claims. This is somewhat less likely, at least within any given prosecutor’s office, where an adverse ruling in a given case may impact other similar cases brought in the future.

However, the most substantial difference between the toxic and the forensic situation is in the ability of admissibility rulings to influence the science itself. On the civil side, lawyers might wish that scientists would do more research of direct relevance to their cases and on rare occasions may even fund

129 “Jean Goldberg, author of a[n]... article on Bendectin and oral clefts, noted, ‘if nothing had been happening over the drug... I doubt even whether I would have written it up.’ JOSEPH SANDERS, BENDECTIN ON TRIAL: A STUDY OF MASS TORT LITIGATION 63 (1998) (quoting MICHAEL D. GREEN, BENDECTIN AND BIRTH DEFECTS: THE CHALLENGE OF MASS TOXIC SUBSTANCES LITIGATION 332 (1996)). Some research may be mandated by the government. The largest animal study on Bendectin was conducted at the specific request of the Food and Drug Administration. JOSEPH SANDERS, BENDECTIN ON TRIAL 63 (1998).


131 Because the parties in some civil cases may have more resources than the parties in a criminal trial, they may be in a better position to respond to an adverse admissibility ruling by seeking out better experts and better existing data in subsequent cases. However, they may not be in a superior position when it comes to generating new data.
research, but by and large the scientists have their own agenda over which the legal system has very little influence. This is not the case with respect to many forensic areas. The legal regime is the primary market for most types of forensic science. Indeed, a substantial percentage of the people who work in this area are the employees of legal entities. This community’s failure to respond to legal signals will have a significant impact on its livelihood. Moreover, the hierarchical structure of the criminal justice system makes it relatively easy for the legal system to organize a research strategy.\footnote{132}

Of course, if there are no scientific improvements to be made, legal influence will come to naught.\footnote{133} But as I noted in the discussion of fingerprint identification,\footnote{134} there are both short run and long run ways to improve the quality of at least some types of forensic evidence.

VII. ADMISSION STRATEGIES FOCUSED ON THE LONGER VIEW

Assuming that I am correct that admissibility decisions can have a greater impact in forensic cases and that using these decisions to push litigators in the direction of better

\footnote{132} I do not want to leave the impression that admissibility decisions have no potential influence on the quality of science in the toxic tort specific causation situation. The path to improving the evidence in these cases is less well defined, but it is not non-existent. Most significantly, the field of toxicogenomics offers the long run possibility that we will be able to ascertain causation at the level of the individual case, radically changing all specific causation testimony. See generally, Gary E. Marchant, Genetics and Toxic Torts, 31 SETON HALL L. REV. 949 (2001); Gary E. Marchant, Genetic Data in Toxic Tort Litigation, 14 J.L. & POL’Y 7 (2006); Jamie A. Grodsky, Genomics and Toxic Torts: Dismantling the Risk-Injury Divide, 59 STAN. L. REV. 1671 (2007).

In the shorter run, genetic information may assist traditional specific causation testimony by allowing experts to make better estimates that the various possible causes of a disease played a role in the plaintiff’s case. See Gary E. Marchant, Genetic Data in Toxic Tort Litigation, 14 J.L. & Pol’y 7 (2006).

In addition, one could envision some level of proficiency testing for specific causation clinical judgments, but the reality is that this approach would be of very limited value because in most cases of interest we simply do not know what caused the plaintiff’s injury and, therefore, we would have no way of knowing if the expert’s judgment was or was not correct.

\footnote{133} For example, we might wish to have some type of proficiency testing in toxic tort cases to ascertain the error rate of physicians making differential diagnoses. Unfortunately, because there is no way to know the ground truth in toxic cases, i.e. what did cause the plaintiff’s illness, proficiency testing is not a viable avenue of investigation.

\footnote{134} See supra Part IV.A.
science is a good thing, how should courts go about this task? There are several alternatives.

Immediately excluding whole areas of testimony. The most radical approach to improving the quality of evidence would be to declare a body of expertise inadmissible until the proponent is able to produce better evidence. The temptation to espouse this solution in the forensics area is fortified by the lackadaisical, even obstructionist attitude of some practitioners in the field. For example, the fingerprint examiner community clings to a set of conceptual perspectives and professional norms that discourage testing. Experts claim that their technique has an error rate of zero. That is, they assert certainty that when they declare a match they have matched the latent print under investigation to the one and only person in the entire world whose finger could have produced the print. This position is the product of what Professors Saks and Koehler label the “Individualization Fallacy,” which they define as “placing an object in a unit category that consists of a single unit. Individualization implies uniqueness.”

\^135 I am far from the first person to argue for this approach. Professor Mnookin makes the same point in the following passage.

Practically speaking, what should Daubert mean? How high a standard of reliability ought trial judges to impose before permitting expert evidence before a jury? One appropriate focal point for the judge, it seems to me, ought to be whether the expert evidence on offer is as reliable as it can reasonably be, considering the particular context and circumstances. Validity under Daubert is not an on/off switch or an all or nothing proposition. It is, at least to some degree, appropriately contextual and gradational: reliability ought to be determined in relation to what information is available or what information could or should have been available with reasonable effort.

Mnookin, supra note 96, at 133; see also Michael J. Risinger, Goodbye To All That, Or a Fool’s Errand, By One of the Fools: How I Stopped Worrying About Court Responses to Handwriting Identification (and “Forensic Science” in General) and Learned to Love Misinterpretations of Kumho Tire v. Carmichael, 43 TULSA L. REV. 447 (2007) (decrying the willingness of some courts to admit every handwriting expert, regardless of the justification for the expert’s position).


\^137 See Simon A. Cole, Out of the Daubert Fire and Into the Fryeing Pan? Self Validation, Meta-Expertise and the Admissibility of Latent Print Evidence in Frye Jurisdictions, 9 MINN. J. L. SCI. & TECH. 453, 470 (2008); Mnookin, supra note 96, at 139. As Mnookin notes, an examiner might face a disciplinary sanction if she were to testify that a match between the latent print and the comparison print was only “likely” or “probable.” Id.

\^138 Michael J. Saks & Jonathan J. Koehler, The Individualization Fallacy in Forensic Science Evidence, 61 VAND. L. REV. 199, 205 (2008). The fallacy leads to statements such as a firearms examiner’s testimony that he is able to identify an unknown weapon “to the exclusion of every other firearm in the world.” United States v. Green, 405 F. Supp. 2d 104, 107 (D. Mass. 2005); see also Paul C. Giannelli, Forensic
position is indefensible.\textsuperscript{139} It engenders a response that says, “if they think that, then they should not be allowed to testify at all.”

Nevertheless, if the history of admissibility decisions in the forensics area teaches us anything, it is that this approach is not going to win favor with the courts. Elected judges might well be committing re-election suicide by excluding key evidence in a criminal case, but even the life-tenured federal judiciary seems unwilling to take such a step.

\textit{Prohibiting testimony about individual causation.} Another option is to limit what experts can say with respect to the particular case. There is some precedent for this solution. When expert evidence on eyewitness identification is permitted, it is generally restricted to general testimony about the factors that affect eyewitness accuracy.\textsuperscript{140} Testimony about whether or not a particular witness is correct is almost always prohibited.\textsuperscript{141} A few courts have placed similar restrictions on

\textit{Science: Under the Microscope,} 34 OHIO N.U. L. REV. 315, 322-24 (2008). Similarly, we have a forensic dentist’s testimony that the defendant “was the only person in the world who could have inflicted the bite marks on [the murder victim’s] body.” Otero v. Warnick, 614 N.W.2d 177, 178 (Mich. App. 2000). Fortunately for the plaintiff in \textit{Otero,} “the Detroit Police Crime Laboratory [later] released a supplemental report that concluded that [he] was excluded as a possible source of DNA obtained from vaginal and rectal swabs taken from [the victim’s] body.” \textit{Id.} at 178. \textit{Otero} is a civil suit by the alleged offender against the dentist for gross negligence in his investigation and testimony. \textit{Id.} The court found that the dentist owed no duty to the plaintiff. \textit{Id.} at 182. For a discussion of the limits of civil suits as a method of holding expert witnesses accountable, see Sanders, \textit{supra} note 123, at 1562-72.

\textsuperscript{139} “Given the general lack of validity testing for fingerprinting, the relative dearth of difficult proficiency tests, the lack of a statistically valid model of fingerprinting and the lack of validated standards for declaring a match, such claims of absolute, certain confidence in identification are unjustified.” Mnookin, \textit{supra} note 96, at 139.

\textsuperscript{140} See 2 FAIGMAN ET AL., \textit{supra} note 35, at § 16.

\textsuperscript{141} Ordinarily, experts on eyewitness identifications—or the courts in which they testify—limit their testimony to the general principles derived from sound empirical research. One exception appears to have been the testimony of Dr. Gerald Loftus in \textit{United States v. Mathis,} 264 F.3d 321 (3d Cir. 2001). Applying the general finding of eyewitness research to the case at hand, according to the court:

\begin{quote}
Dr. Loftus stated that “it’s two to three times as likely that the identification in the photo montage was made based on seeing the photograph four weeks earlier than it was based on seeing the individual” who fled on October 14, 1998.
\end{quote}

\textit{Id.} at 334.

At least one treatise written by a well known investigator in the field made the following comments concerning this type of testimony:

\begin{quote}
Without the transcript of the trial, it is impossible to say what literature Dr. Loftus relied upon in making these conclusions regarding the likely accuracy of a particular witness’ specific identification. To our knowledge, there is no research support in the scientific literature for the validity of such clinical opinions. In psychology, work on eyewitness identifications has, on the whole,
handwriting experts. For example, the expert is permitted to testify about general features that distinguish individual handwriting, but not whether a particular signature is a forgery.\footnote{Paul Giannelli, supra note 138, at 319 n.30 notes that this was the position of the courts in United States v. Rutherford, 104 F. Supp. 2d 1190, 1194 (D. Neb. 2000) and United States v. Hines, 55 F. Supp. 2d 62, 67 (D. Mass. 1999).}

As attractive as this alternative may be in some criminal law contexts, were it ever to be used on the civil side it would run directly into the plaintiff's burden to prove specific causation by a preponderance of the evidence. And simply because there is no expert opinion on the individual causation question does not mean that the jury can avoid the issue. It must decide if this signature is a forgery or if the plaintiff's disease was caused by the defendant. We do not know enough about jury decision making to know whether the quality of the jury's judgment would benefit from restricting expert testimony in this way.

\textit{Exclude the worst testimony.} This alternative assumes that even in areas where there is little science we may still distinguish levels of reliability, and that some testimony is so unreliable that it fails a minimal threshold of reliability. One way to think about such a situation is to conclude that the expert is not an expert at all because the area of so-called expertise does not exist. The "expert" witness is a lay witness in disguise.

This approach has the advantage of incrementalism. One does not need to make global pronouncements concerning a particular type of expert testimony. For example, courts might decide to admit handwriting expert testimony when the task is to determine whether a signature is a forgery based on many exemplars of known authentic signatures while excluding testimony asserting the authorship of an attempted forgery.\footnote{See generally Risinger, supra note 135.}

---

\textit{2 Faigman, et al., supra note 35, at § 16:2.}
The approach requires the courts to draw difficult lines and thus it threatens to produce inconsistent outcomes in close cases. Nevertheless, this may be the best that we can do in areas where admissibility decisions will not have a foreseeable impact on future cases or future science and the only goal courts can hope to achieve is to change the mix of cases by eliminating the weakest among them.

*Excluding evidence at some point in the future.* There is an additional alternative that directly focuses on the longer view. A court could declare that at some future time the proponent of a type of evidence must present a better justification for its reliability or risk the exclusion. A precise timetable is not required. The court does not need to establish a date certain when evidence will no longer be admitted without better science. A vaguer nod to the future may well suffice. For example, the court could tell the United States that it should ask the F.B.I. to engage in significant, serious proficiency testing.\(^\text{144}\)

For all cases coming before district courts two or three years hence, experts testifying about a fingerprint identification should be prepared to present the results of such tests or indicate why this is not yet possible. The details of such an opinion are less important than the general requirement that the proponent work toward better justification for its position.

I know of no cases that have adopted this position. In most situations, because of the very limited impact legal opinions have on advancements in our causal understanding, it would be unnecessary, inappropriate, or both. But in the forensic area we find ourselves in a unique chicken-and-egg situation. Because courts are committed to a contextual approach to justification, they permit forensic expert testimony with very little warrant. But this liberal approach unnecessarily helps to perpetuate the status quo. Absent the very unlikely prospect of a congressional mandate, the courts are in the best position to move us toward a time when forensic evidence stands on a more solid scientific foundation.

\(^{144}\) Because of the F.B.I.'s central role in the production of forensic evidence, the greatest impact would come from a federal circuit court opinion sets a timetable for improved evidence. One must be wary, of course, of the quality of research generated by the very organization being asked to provide evidence. On this point, the F.B.I.'s track record is anything but reassuring. For an insightful article on the less than evenhanded nature of past law-enforcement sponsored research efforts, see D. Risinger & Saks, *supra* note 112, at 1023. Given this record, a court might wish to stipulate in advance that the research is undertaken by an independent entity.
VIII. SUMMARY

Following the United States Supreme Court’s opinion in *Daubert v. Merrell Dow Pharmaceuticals*, trial courts are increasingly asked to assess the reliability of expert testimony. Testimony that is deemed to be insufficiently reliable is excluded. Many scholars argue that the courts apply the reliability requirement inconsistently. One frequently mentioned example is the use of liberal admissibility standards in criminal law forensic evidence cases and the use of stringent admissibility standards in tort law toxic tort cases. I argue that this apparent inconsistency misses a key point.

Admissibility decisions in toxic tort cases appear to be stringent because they address two separate causal questions: evidence of general causation and evidence of specific causation. When we focus on admissibility decisions dealing exclusively with specific causation we see that courts adopt a liberal standard. I compare this standard with liberal admissibility standards in forensic cases, using fingerprint evidence as an example.

In both situations, courts must decide a causal question with respect to a particular individual and often the only expertise available to address the issue is the judgment of a clinician, be that clinician a medical doctor or a fingerprint analyst. In neither situation does the expert have much by way of quantified, empirical evidence to support her conclusion. It is the relative lack of scientific information that causes the courts to lower their admissibility standard in both these situations.

This liberal admissibility standard is consistent with law’s contextual approach to knowledge. A contextual approach varies the amount of justification necessary for an expert to express an opinion depending on the quantity and quality of evidence on the point. Admissibility standards with respect to questions of individual causation are relatively liberal because we have relatively little systematic scientific evidence on point.

Unfortunately, the contextual approach does not create any incentive for the parties to improve the quality of evidence even in situations where this might be possible. I argue that some branches of forensics are such situations. When this occurs, courts should view the relevant context from a longitudinal perspective and adjust their admissibility rulings accordingly.
Religious Accommodation and Housing

FAIR HOUSING AFTER BLOCH V. FRISCHHOLZ

Proper enforcement of the Fair Housing Act’s promise of equal housing opportunity and of the First Amendment’s guarantee to protect the practice of religion without the government establishing religion can help ensure that all persons live comfortably together in our pluralistic society and that all persons have access to safe, decent, sanitary housing where they can exercise their right to worship or not to worship as they choose.¹

INTRODUCTION

As this quote astutely recognizes, the synergy of the distinct goals of the Fair Housing Act² and the Establishment Clause of the First Amendment³ has the ability to produce an ideal situation—a nation where individuals of all religions have access to adequate housing in a place where they are free to exercise their religious beliefs. This utopia, however, has yet to be achieved. More than forty years after the passing of the Fair Housing Act, major barriers still exist and continue to be created, preventing individuals from living in places where they are free to act in accord with their religious beliefs. In fact, as evidenced by Bloch v. Frischholz,⁴ where the Seventh Circuit initially upheld a condominium association’s rule prohibiting

³ The First Amendment’s Establishment Clause states, “Congress shall make no law respecting an establishment of religion.” U.S. Const. amend. I. The religion clauses of the First Amendment are applicable to the states via incorporation by the Fourteenth Amendment. Cantwell v. Connecticut, 310 U.S. 296, 303 (1940).
⁴ Bloch v. Frischholz (Bloch I), 533 F.3d 562 (7th Cir. 2008), aff’d in part, rev’d in part, 587 F.3d 771 (7th Cir. 2009) (en banc).
the display of anything, including religiously-mandated objects, in the doorways of the condominiums, judicial decisions have the power to further the rift between the allied goals of the First Amendment and the Fair Housing Act, namely providing equal housing opportunities and religious freedom to all Americans. Decisions such as these can, and in some cases have, made it harder for individuals in protected classes to obtain access to housing where they are free from religious persecution. Today, some individuals in the United States are faced with the decision of compromising their religious beliefs or moving out of their homes.

In 1968, the Fair Housing Act was enacted as an effort to control the pervasive discrimination in the housing market. Prohibiting discrimination by both public and private housing providers, the statute lays out protected classes of individuals, including the religiously observant, who may not be subject to discrimination. Because the Fair Housing Act applies to both public and private housing, it creates First Amendment obligations for private entities that previously did not exist. The First Amendment’s Free Exercise and Establishment Clauses prohibit laws that burden the practice of religion and government action which promotes religion, respectively. As such, housing providers are required to provide non-discriminatory housing in a way that neither favors nor disadvantages the free exercise of religion.

On July 10, 2008, the Seventh Circuit affirmed the constitutionality of a condominium association’s rule that effectively achieves the opposite of the utopian dream described

---

5 See Boodram v. Md. Farms Condo., No. 93-1320, 1994 WL 31025, at *1 (4th Cir. Feb. 1, 1994) (affirming summary judgment in favor of defendant condominium association despite plaintiff’s claim that the association’s rule prohibiting storage on condominium balconies interfered with his religious duty to display red flags, known as “Jhandee,” as compelled by the Hindu faith); see also Savanna Club Worship Serv. v. Savanna Club Homeowners’ Ass’n, 456 F. Supp. 2d 1223, 1224, 1234 (S.D. Fla. 2005) (holding that homeowners’ associations have the right to prohibit religious worship in the common areas of their communities).
6 Seng, supra note 1, at 1; see also U.S. Dep’t. of Housing and Urban Dev., History of Fair Housing, http://www.hud.gov/offices/fheo/aboutfheo/history.cfm (last visited Aug. 30, 2009).
7 See 42 U.S.C. § 3604(a) (stating that it is unlawful “[t]o refuse to sell or rent after the making of a bona fide offer, or to refuse to negotiate for the sale or rental of, or otherwise make unavailable or deny, a dwelling to any person because of race, color, religion, sex, familial status, or national origin.”).
8 U.S. CONST. amend. I.
9 Seng, supra note 1, at 1.
In its initial review of *Bloch v. Frischholz* (*Bloch I*), a two-to-one decision, the court affirmed the Northern District of Illinois’ grant of summary judgment in favor of a condominium association whose prohibition on the placement of anything in the doorways of the condominiums, including religiously-mandated objects, was challenged by one of the condominium owners. The decision essentially held that condominium associations may make rules that inhibit the exercise of religion so long as the rule is “neutral and applicable.” In particular, the court found that even though observant Jewish condominium owners felt prohibited from living in their homes once the condominium rule was interpreted to prohibit the display of *mezuzot,* the rule was not in violation of either the First Amendment or the Fair Housing Act, which, as currently written, does not require accommodation for religion. On November 13, 2009, six months after the Seventh Circuit reheard the case en banc, the court decided that although the condominium’s rule did not make the condominiums “unavailable,” factual issues did exist with regard to the issue of intentional discrimination, rendering total summary judgment improper and remanding the case for further proceedings.

This Note argues that the Seventh Circuit’s en banc ruling in *Bloch v. Frischholz* (*Bloch II*) only partially remedies the potentially harsh policy implications that could have resulted from the circuit’s initial ruling. The en banc decision failed to recognize that the actions taken by the condominium

---

10 See *Bloch v. Frischholz* (*Bloch I*), 533 F.3d 562, 565 (7th Cir. 2008), aff’d in part, rev’d in part, 587 F.3d 771 (7th Cir. 2009) (en banc).
11 Id. at 563. Judges Easterbrook and Bauer were in the majority and Judge Wood wrote in dissent.
12 Id. at 566 (“Plaintiffs would like us to treat failure to make an accommodation as a form of discrimination. That was one theme of Justice O’Connor’s separate opinion in *Smith*—but the majority held that a neutral, exception-free rule is not discriminatory and is compatible with the Constitution’s free exercise clause.”).
14 *Bloch I*, 533 F.3d at 565.
15 *Bloch v Frischholz* (*Bloch II*), 587 F.3d 771 (7th Cir. 2009).
board rendered the units unavailable to observant Jews.\textsuperscript{16} Thus, the new ruling does not completely nullify the ability of housing associations to create rules like this in the future. Furthermore, if on remand the issue of intentional discrimination is not resolved in the Blochs’ favor, the door will remain open for discriminatory housing practices to prevail. Finally, because this decision is not binding on all federal courts, courts outside the jurisdiction of the Seventh Circuit could easily render the same mistaken ruling as the \textit{Bloch I} court, which would result in consequences inconsistent with both the congressional intent of the Fair Housing Act and the general common law on issues of religious accommodation in contexts parallel to housing. In order to prevent judicial decisions such as these from producing results adverse to those envisioned by the legislature, it is vital that the legislature take action to avert the disparaging consequences that would otherwise result.

Part I of this Note discusses the factual underpinnings of the \textit{Bloch} decisions to give context to the Seventh Circuit’s decision and the analysis contained herein. Part II briefly discusses the historical interaction between the Fair Housing Act and the First Amendment, focusing on how these laws work together and affect one another. Part III analyzes the legislative history and congressional intent of the Fair Housing Act and examines how religious accommodation has been dealt with in the employment context, where there is a statutorily imposed religious accommodation requirement. This analysis seeks to assess whether \textit{Bloch I} and \textit{II} were rightfully decided and explore the potential for a religious accommodation requirement under the Fair Housing Act in the future. Finally, Part IV discusses the implications of the \textit{Bloch} decisions and suggests a road for moving forward. These suggestions aim to provide a way to overcome the harsh impact that such decisions could produce. In light of the disturbing potential consequences of the initial decision, and the fact that these consequences have not been completely obviated, it is imperative that Congress act to rectify the judiciary’s failure to protect the spirit of the Fair Housing Act.

\textsuperscript{16} \textit{Id.} at 776-84.
I. THE CASE: BLOCH V. FRISCHHOLZ

In a two-to-one decision in Bloch v. Frischholz, upon first review the Seventh Circuit held that a condominium association’s rule of not allowing anything to be placed in condominium doorways, including religiously-mandated objects, did not violate either the First Amendment or the Fair Housing Act, and was therefore constitutional.\textsuperscript{17} Finding the rule to be neutrally applicable to all condominium owners regardless of their religious beliefs, the court affirmed the district court’s grant of summary judgment in favor of the condominium association.\textsuperscript{18} This decision is incongruous with the legislative intent of the Fair Housing Act and the established legal precedent in parallel contexts. These incongruities were not completely remedied by the Seventh Circuit’s en banc rehearing, where the court found that although issues of fact existed as to the Blochs’ claim of intentional discrimination, the condominium association’s rule was not in violation of the Fair Housing Act.\textsuperscript{19}

A. Bloch v. Frischholz: The Facts

Lynne, Helen, and Nathan Bloch, observant Jews and residents of the Shoreline Towers condominium building in Chicago, Illinois, brought suit against the Shoreline Towers Condominium Association and its president, Edward Frischholz, alleging intentional discrimination in violation of the Fair Housing Act\textsuperscript{20} and the Civil Rights Act.\textsuperscript{21} The claims arose out of a “hallway rule” that the condominium association adopted in September 2001, while Lynne Bloch was the chair of the rules committee. The rule stated: “1. Mats, boots, shoes, carts, or objects of any sort are prohibited outside Unit entrance doors. 2. Signs or name plates must not be placed on Unit doors.”\textsuperscript{22} The rule did not become problematic until 2004, when the association began to interpret the rule as prohibiting

\textsuperscript{17} Bloch I, 533 F.3d at 565.
\textsuperscript{18} Id. at 564-65.
\textsuperscript{19} Bloch II, 587 F.3d 771.
\textsuperscript{20} 42 U.S.C. §§ 3604(a)-(b), 3617 (2006).
\textsuperscript{21} Id. § 1982; Bloch I, 533 F.3d at 569 (Wood, J. dissenting).
\textsuperscript{22} Bloch I, 533 F.3d at 567 (Wood, J. dissenting).
the placement of *mezuzot* on the exterior doorposts of the units.\(^{23}\)

During a renovation of the hallways of the Shoreline Towers in May 2004, the residents were instructed to remove all items from their doors so that construction could be completed.\(^{24}\) The Blochs complied, and once the renovations were completed, proceeded to reattach their *mezuzah* to the entrance of their unit.\(^{25}\) Shortly thereafter, the condominium association began removing and confiscating *mezuzot* from entranceways in the building, claiming that they were in violation of the hallway rule.\(^{26}\) Previously, the rule had only been used to prevent clutter in the hallway as well as “signs and name plates” as explicitly stated, but not *mezuzot*.\(^{27}\) However, despite plaintiffs’ objections, explaining the religious significance and importance of the *mezuzah*, the condominium association offered no relief. Instead, the condominium association continually removed plaintiffs’ *mezuzot* and threatened monetary penalties if they continued to affix a *mezuzah* in their doorway.\(^{28}\)

After the death of Marvin Bloch, Lynne’s husband and Helen and Nathan’s father, the Blochs specifically requested permission to display a *mezuzah* in accordance with Jewish mourning rituals.\(^{29}\) Despite their request, the condominium association removed their *mezuzah* during this traumatic time.\(^{30}\) Debra Glassman, another Shoreline Towers resident and observant Jew, was treated in a similar fashion, forcing her to move out of her unit.\(^{31}\) She felt that “she had essentially been evicted from her home,” because the condominium association prevented her from displaying her *mezuzah* as required by the laws of her religion.\(^{32}\)

\(^{23}\) Id.

\(^{24}\) Id.

\(^{25}\) Id.

\(^{26}\) Id.

\(^{27}\) Id.

\(^{28}\) Id.

\(^{29}\) Id.

\(^{30}\) Id.

\(^{31}\) Id.

\(^{32}\) Id.; see also Deuteronomy 6:5-9 (instructing all Jews that “[y]ou shall write [the commandments] on the doorposts of your house and on your gates”). Both orthodox and non-orthodox Jews affix the prayer to their doorpost, as required, in a *mezuzah*. Dennis W. Carlton & Avi Weiss, *The Economics of Religion, Jewish Survival, and Jewish Attitudes Toward Competition in Torah Education*, 30 J. LEGAL STUD. 253, 263 n.27 (2001). *Mezuzot* are required because they serve as a constant reminder of the
In the aftermath, Lynne Bloch proposed an amendment to the hallway rule that would allow mezuzot to be displayed on exterior doorways, to no avail. In December 2006, the Blochs filed suit in the Northern District of Illinois. Before the Blochs filed their suit, however, the Shoreline Towers Condominium Association’s board adopted a religious exception to the hallway rules. Thus, the relief sought by the Blochs merely consisted of damages for the distress suffered before the exception was enacted, as well as an injunction to prevent the association from restoring the old interpretation of the rule. Additionally, about a year before the Blochs filed their complaint, the city of Chicago enacted an ordinance prohibiting residential building owners from restricting the placement of religious objects in the doorways of homes, unless necessary to avoid property damage or undue hardship to other unit owners. This ordinance made it illegal for the Association to revert to its prior version of the rule. Thus, the Blochs’ suit essentially only involved a quest for damages, allowing the issue raised by the Blochs to remain ripe for adjudication.

B. Bloch v. Frischholz: The Suit

The district court granted summary judgment in favor of the condominium association and Frischholz based on the Seventh Circuit’s holding in Halprin v. Prairie Single Family Homes of Dearborn Park Association, and the Seventh Circuit affirmed. Writing for the circuit court, Chief Judge Frank Easterbrook explained that Halprin stood for the proposition that harassment of owners or tenants, even though religiously motivated, was not a Fair Housing Act violation. Only when...
the harassment is so severe as to amount to constructive eviction can a violation be found.\textsuperscript{42} Finding the hallway rule neutral with respect to religion because it was generally applicable to all condominium owners, the court held that a determination of whether the rule resulted in the constructive eviction of the Blochs was unnecessary.\textsuperscript{43} Finally, the court explained that it saw the suit as seeking a religious exception to a neutral rule—a religious accommodation—something not required by the language of the Fair Housing Act.\textsuperscript{44} Because the Fair Housing Act sees discrimination as more than a “failure to accommodate,” the court found no violation of the Fair Housing Act or the First Amendment in Shoreline Towers’ refusal to accommodate the religious beliefs of the Jewish homeowners.\textsuperscript{45}

In a vehement dissent, Judge Diane Wood disagreed with the majority’s grant of summary judgment.\textsuperscript{46} She noted that the court wrongly framed the issue on appeal and thus did not properly address the claims asserted by the plaintiffs.\textsuperscript{47} According to Judge Wood, instead of viewing the Blochs’ claim as a quest for religious accommodation, the court should have reached the question of whether the inability to display a mezuzah in one’s doorway resulted in the constructive eviction of observant Jewish residents, and should have found that plaintiffs produced sufficient evidence to prove constructive eviction.\textsuperscript{48} Judge Wood found the hallway rule to be the equivalent of a sign outside the building reading “No observant Jews allowed,” and as such, the situation presented in the case was exactly that imagined by Fair Housing Act’s Section 3604(a) as interpreted in \textit{Halprin}.\textsuperscript{49} Furthermore, Judge Wood

\begin{itemize}
    \item \textsuperscript{42} Id. at 564.
    \item \textsuperscript{43} Id.
    \item \textsuperscript{44} Id. at 565 (explaining that the language of the Fair Housing Act only requires accommodations for handicaps and not for religion or the other classes protected by the Act).
    \item \textsuperscript{45} Id.
    \item \textsuperscript{46} Id. at 566.
    \item \textsuperscript{47} Id. (The Court addressed the issue of whether the condominium association had a duty to accommodate plaintiffs’ religious beliefs rather than analyzing the actual issue, that is, whether the association intentionally discriminated against plaintiffs based on their religious beliefs.).
    \item \textsuperscript{48} Id. at 570.
    \item \textsuperscript{49} Id. In \textit{Halprin}, the Seventh Circuit found that religiously motivated harassment of owners and tenants did not to violate the Fair Housing Act after it gave a limited interpretation of Section 3604(b) and refused to look beyond the plain meaning of the words contained in the statute. See \textit{Bloch}, 533 F.3d at 563; see also \textit{Halprin v. Prairie Single Family Homes of Dearborn Park Ass’n}, 388 F.3d 327 (7th Cir. 2004) (affirming summary judgment in favor of defendant homeowners’ association,}
noted that the Section 3604(b) claim asserted by the Blochs was factually sufficient because the Fair Housing Act prohibited, among other things, religious discrimination in “the terms, conditions, or privileges of sale or rental of a dwelling,” and the agency responsible for regulating the implementation of the Fair Housing Act, the United States Department of Housing and Urban Development (“HUD”), had interpreted this language to protect “an owner, tenant, or a person associated with him or her.”

Judge Wood argued that summary judgment was incorrect because the question of whether the hallway rule applied to mezuzot was both material and in dispute. Thus, the majority’s characterization of the rule as being facially neutral was improper, since that would require a finding that the rule does include a prohibition on the display of mezuzot. According to Judge Wood, the whole point of the case is that the Association took a neutral rule and started interpreting it in a way that exclusively affected observant Jewish owners. In such a situation, it is not just the fact that a rule is neutral that is of importance. Rather, she argued that it is necessary to assess whether the rule “target[s] the practices of a particular religion for discriminatory treatment” in order to determine if the Free Exercise Clause of the First Amendment has been violated. In Judge Wood’s opinion, this was exactly the situation at bar—the rule may have been facially neutral, but it had a disparate impact on observant Jews, therefore invaliding its neutrality.

Finally, Judge Wood opined that the Blochs’ claim under 42 U.S.C. § 1982 was also sustainable. and finding that religious harassment of plaintiff homeowners by association president did not violate § 3604(b) because it was not harassment that “prevented [people] from acquiring property” since the couple already owned their home).
On May 13, 2009, the Seventh Circuit reheard the Blochs’ case en banc.\(^57\) The United States government felt so strongly that the Seventh Circuit erred in its initial ruling that it submitted an amicus brief urging the en banc panel to reverse and remand the case.\(^58\) In its brief, the government argued that the Seventh Circuit’s ruling is out of line with the congressional intent of the Fair Housing Act and that summary judgment was inappropriate because enough questions of fact existed to make the case ripe for jury determination as to the existence of discrimination.\(^59\) The government conceded that the text of the Fair Housing Act does not currently require religious accommodation, but remained silent on the issue of whether such a clause is inferred by the spirit of the Act or the Constitution.\(^60\)

Six months after hearing the case en banc, the Seventh Circuit partially amended its initial ruling.\(^61\) The en banc court saw the case as presenting two distinct issues. The first issue was which, if any, of the Fair Housing Act provisions could be a potential source of relief for the Blochs.\(^62\) The second issue was whether the Blochs put forth sufficient evidence of discrimination to create an issue of fact to be resolved at trial.\(^63\) As to the first issue, the court looked at whether the Fair Housing Act can afford relief for claims of post-sale discrimination.\(^64\) While the court acknowledged that Section 3604(a) can be violated post-acquisition in extreme cases of “constructive eviction,” it found that the condominium at issue here was never actually made “unavailable” to the Blochs, and therefore affirmed the grant of summary judgment on this issue.\(^65\) On the issue of intentional discrimination, the court

\(^{57}\) Bloch v. Frischholz (Bloch II), 587 F.3d 771 (7th Cir. 2009). For an audio version of the argument, see 2009 WL 1472344.

\(^{58}\) Brief for the United States as Amicus Curiae in Support of Plaintiffs-Appellants Urging Reversal and Remand on Fair Housing Act Claims, Bloch v. Frischholz, 533 F.3d 562 (7th Cir. 2008) (No. 06-3376).

\(^{59}\) Id. at *13-14.

\(^{60}\) Id. at *43-44. This Note seeks to argue that although the Fair Housing Act does not contain a religious accommodation clause, such a clause is necessary for the Act to be congruent with the congressional intent of the statute, the First Amendment, and how religious accommodation is handled in parallel contexts, namely employment under Title VI. See infra Part III.A-C.

\(^{61}\) Bloch II, 587 F.3d at 787.

\(^{62}\) Id. at 775.

\(^{63}\) Id. at 775-76.

\(^{64}\) Id. at 775.

\(^{65}\) Id. at 777-78 (“To establish a claim for constructive eviction, a tenant need not move out the minute the landlord’s conduct begins to render the dwelling
found that the Blochs had presented sufficient evidence of genuine issues of fact to warrant a trial. Consequently, the district court’s grant of summary judgment and the circuit panel’s affirmation was reversed, and the Blochs’ case was remanded.

II. THE FAIR HOUSING ACT AND THE FIRST AMENDMENT

Title VIII of the Civil Rights Act, or the Fair Housing Act, was enacted in 1968 during the height of the Civil Rights Movement, and in the aftermath of the assassination of Martin Luther King, Jr., as a congressional effort to curb rampant discrimination in the housing market. The Act prohibits discrimination in the sale, rental, and housing provisions of both public and private housing providers, subject to certain exemptions, on the basis of “race, color, religion, sex, familial status, [and] national origin.” Through the inclusion of religion in the list of protected classes, the drafters implicated the First Amendment in the Fair Housing Act’s interpretation. Although the First Amendment generally applies only to public entities, because the Fair Housing Act outlaws discrimination in both private and public housing, it effectively creates First Amendment obligations for private housing authorities. Therefore, in applying the Fair Housing Act, a delicate balance must be struck between neither favoring nor disadvantaging religion, as mandated by the First Amendment.

uninhabitable—in this case, when the defendants began enforcing the Hallway Rule to take down the Blochs’ mezuzot. Tenants have a reasonable time to vacate the premises. Nonetheless, it is well-understood that constructive eviction requires surrender of possession by the tenant. Still, the Blochs never moved out.” (internal citations omitted)).

66 Id. at 785 (“Although the Blochs’ case is no slam dunk, we think the record contains sufficient evidence, with reasonable inferences drawn in the Blochs’ favor, that there are genuine issues for trial on intentional discrimination.”).

67 Id. at 787.


69 Seng, supra note 1, at 1; see also U.S. Dep’t. of Housing and Urban Dev., History of Fair Housing, http://www.hud.gov/offices/fheo/abouttheo/history.cfm (last visited Aug. 30, 2009).

70 See 42 U.S.C. § 3603(b).

71 Id. § 3604(a)-(b).

72 See, e.g., id. § 3604(a) (stating that it is unlawful to “refuse to sell or rent after the making of a bona fide offer, or to refuse to negotiate for the sale or rental of, or otherwise make unavailable or deny, a dwelling to any person because of race, color, religion, sex, familial status, or national origin”).

73 Seng, supra note 1, at 1.
Congress intended the Fair Housing Act to “provide, within constitutional limitations, for fair housing throughout the United States,”74 in an effort to promote racially integrated housing.75 Although primarily designed as a remedial tool to deal with segregated housing patterns in the United States,76 the statute not only sought fair housing for individuals discriminated against on the basis of race, but also took on the broader task of providing fair housing to other individuals who are likely to be the subject of discrimination.77 Both the Department of Justice, by way of the Attorney General, and individuals are eligible to bring suit under the Fair Housing Act.78 However, while individuals are only required to show that they have been victims to an illegal housing practice, the Department of Justice must point to a “pattern or practice” of discrimination in order to establish a cause of action.79

The Supreme Court has had surprisingly minimal interaction with substantive Fair Housing Act claims in the forty years since its enactment.80 Circuit courts, however, have had extensive engagement with Fair Housing claims and have developed standards under which evaluation of such claims are to be assessed. Specifically, these courts have determined that violations of the Fair Housing Act can be established on one of two grounds—disparate impact or disparate treatment.81

Generally, to establish a disparate impact claim, a claimant must make a prima facie showing of discrimination, demonstrating that the defendant’s conduct has a discriminatory effect.82 More specifically, a claimant must prove that the challenged practice “actually or predictably” results in

74 42 U.S.C. § 3601.
75 Otero v. N.Y. City Hous. Auth., 484 F.2d 1122, 1134 (2d Cir. 1973).
77 See 42 U.S.C. § 3604 (prohibiting discrimination on the basis of “race, color, religion, sex, familial status, or national origin”).
78 See id. §§ 3613-14.
79 The Fair Housing Act, supra note 76.
80 Supreme Court cases on the Fair Housing Act have mostly been limited to addressing procedural issues such as standing under the Fair Housing Act. See Havens Realty Corp. v. Coleman, 455 U.S. 363, 366 (1981); Gladstone, Realtors v. Bellwood, 441 U.S. 91, 93 (1978); Trafficante v. Metro. Life Ins. Co., 409 U.S. 205, 208 (1972).
81 Leblanc-Sternberg v. Fletcher, 67 F.3d 412, 425 (2d Cir. 1995).
82 This is similar to the requirement for alleging a cause of action under Title VII of the Civil Rights Act. See Williams v. Matthews Co., 499 F.2d 819, 826 (8th Cir. 1974); see also infra Part III.C.
discrimination based on one of the prohibited classifications.\footnote{United States v. City of Black Jack, Mo., 508 F.2d 1179, 1184-85 (8th Cir. 1975); see also United Farmworkers of Fla. Hous. Project, Inc. v. City of Delray Beach, 493 F.2d 799, 808 (5th Cir. 1974).} A claimant need not show that such action is discriminatorily motivated, but merely that the defendant’s action has a discriminatory effect.\footnote{City of Black Jack, Mo., 508 F.2d at 1184-85.} For these types of claims, a claimant is required to prove a “causal connection” between the questionable policy and the resulting disparate impact on the protected group.\footnote{Brown v. Coach Stores, Inc., 163 F.3d 706, 712 (2d Cir. 1998) (citing Lopez v. Metro. Life Ins. Co., 930 F.2d 157, 160 (2d Cir. 1991)).}

Alternatively, Fair Housing Act violations can be established based on a disparate treatment theory. Under this theory, a claimant can establish a prima facie showing of discrimination by “showing that animus against the protected group ‘was a significant factor in the position taken.’”\footnote{Leblanc-Sternberg, 67 F.3d at 425 (quoting United States v. Yonkers Bd. of Educ., 837 F.2d 1181, 1226 (2d Cir. 1987)).} Allegations of discriminatory intent must be analyzed based on the totality of the circumstances, taking into account the fact that some “law[s] bear more heavily” on one group of people than others.\footnote{Washington v. Davis, 426 U.S. 229, 242 (1976).} Other factors considered in this analysis are the historical background of the decision, the events leading up to the decision, and statements made by individuals involved in the decision-making.\footnote{Vill. of Arlington Heights v. Metro. Hous. Dev. Corp., 429 U.S. 252, 267-68 (1977).}

The type of evidence used to make out a claim of religious discrimination under the Fair Housing Act is the same type used in assessing the religious animus of a law for the purpose of assessing its constitutionality under the First Amendment.\footnote{Leblanc-Sternberg, 67 F.3d at 426 (citing Church of Lukumi Babalu Aye, Inc. v. City of Hialeah, 508 U.S. 520, 540 (1993)).} Specifically, this type of evidence is used to prove religious animus for disparate treatment claims. See id.

Under the First Amendment, in order for a facially neutral rule to constitutionally prohibit conduct that inhibits the practice of religion, there must be a religiously
neutral reason to justify the rule.\textsuperscript{91} Thus, claims of religious discrimination brought under the Fair Housing Act necessarily implicate the right to the free exercise of religion guaranteed by First Amendment.\textsuperscript{92}

III. THE SEVENTH CIRCUIT’S MISSTEP: ANALYSIS OF BLOCH

The Seventh Circuit’s decisions in Bloch I and II were wrongly decided for a number of reasons. To begin, the court’s consistent conclusion that the Fair Housing Act does not require accommodation for religion fails to look beyond the words of the statute in assessing its applicability to the situation faced by the Blochs. In addition, the Bloch I panel’s classification of the hallway rule in question as neutrally-applicable, and thus consistent with the requirements of the First Amendment, fails to consider disparate impact analysis.\textsuperscript{93} Furthermore, the decision is inconsistent with how religious accommodation has been dealt with in the parallel context of workplace discrimination. Finally, the policy implications of the initial holding are problematic, and have vast consequences for the future of the housing market that the en banc ruling did not ameliorate. Depending on how the case is decided on remand and whether other circuits follow the rehearing of the Seventh Circuit’s initial ruling, the private housing market may now be able to effectively exclude protected classes, making the Fair Housing Act’s goal of rendering adequate housing to all individuals, regardless of their religious beliefs, unachievable.

Part A of this Section discusses why the congressional intent of the Fair Housing Act requires religious accommodation, despite the absence of explicit language to this effect. Part B analyzes the need for a religious accommodation clause under the Fair Housing Act because of the First Amendment implications in the absence of such a requirement. Next, Part C assesses how religious accommodation is dealt with under Title VII of the Civil Rights Act, where religious accommodation is statutorily mandated, in an effort to prove that such a requirement is consistent with First Amendment jurisprudence in the employment context. Finally, Part D looks

\textsuperscript{91} See Church of Lukumi Babalu, 508 U.S. at 531-32.
\textsuperscript{92} U.S. CONST. amend. I.
\textsuperscript{93} See Bloch v. Frischholz (Bloch I), 533 F.3d 562, 573 (7th Cir. 2008) (Wood, J., dissenting), aff’d in part, rev’d in part, 587 F.3d 771 (7th Cir. 2009) (en banc).
at disparate impact analysis to show why the Seventh Circuit twice erred in its failure to assess the Blochs’ claim on these grounds.

A. Reconciling Bloch and the Congressional Intent of the Fair Housing Act

The Fair Housing Act does not contain an explicit requirement that public or private housing providers make accommodations for religious individuals. But looking beyond the plain meaning of the words contained in the Act—to the congressional intent in enacting this legislation—reveals that such a requirement is essential in order for the Act to achieve its stated goals. Expanding the scope of statutory interpretation afforded to the Fair Housing Act by the Seventh Circuit exposes the fact that its narrow reading fails to give proper breadth to the Act, and consequently fails to grant relief to those individuals who are harmed by housing providers’ failure to make reasonable accommodations for the religiously observant. These results are in direct conflict with the spirit of the Act, and its ability to “provide . . . for fair housing throughout the United States.” Indeed, the Supreme Court has itself endorsed a liberal reading of the Act, providing more evidence that the Seventh Circuit’s narrow interpretation was mistaken.

When the Fair Housing Act was enacted in 1968, it only prohibited discrimination in the “sale, rental and financing” of housing on the basis of “race, color, religion, sex or national origin.” In 1988, the Fair Housing Act was amended to increase the number of protected classes and to provide better guidelines for enforcement of the statute and the rights provided therein. The amendments extended the guarantees of the statute to cover individuals with disabilities and to protect against discrimination on the basis of familial status. Initially, the statute contained no mention of the concept of “accommodation” for any of the protected classes, but with the

---

94 See 42 U.S.C. §§ 3601-3619; see also Bloch I, 533 F.3d at 565; Hack v. President of Yale Coll., 237 F.3d 81, 88 (2d Cir. 2000).
95 See supra note 75 and accompanying text.
97 See infra note 106 and accompanying text.
98 The Fair Housing Act, supra note 76.
99 Id.
100 Id.
1988 amendment, a clause was added that stated that “a refusal to make reasonable accommodations in rules, policies, practices, or services, when such accommodations may be necessary to afford such person equal opportunity to use and enjoy a dwelling,” constituted discrimination against individuals on the basis of a handicap. Accordingly, accommodations for handicapped individuals must be made where the modification is necessary for the individual to use and fully enjoy the dwelling and the modification does not pose an undue cost or administrative burden.

Although the accommodation clause of the Fair Housing Act appears only under the requirements for providing housing to individuals with handicaps, whether this distinction was intended is unclear. Specifically, it is unclear whether Congress intended for the accommodation language to apply retroactively to all forms of discrimination prohibited by the statute. In fact, there are indications that this intent was present, since to provide otherwise would inhibit the stated purpose of the statute, “to provide . . . for fair housing throughout the United States.”

In *Bloch I*, the Seventh Circuit indicated that because the word “accommodate” only appears under the requirements for handicaps, Congress intended the word “discriminate” to have a distinctly different meaning than “failure to accommodate.” However, although the Seventh Circuit is correct that the statutory language does not equate “discrimination” and “failure to accommodate,” and that not all failures to accommodate would rise to the level of discrimination, there may be instances where a failure to accommodate does reach the level of discrimination. Thus, the question of whether an accommodation requirement does, or should, exist for the protected class of religious individuals becomes central to analyzing whether the conduct of the Shoreline Towers Condominium Association rose to the level of

---

102 The cost of the accommodation must also be paid for by the individual requiring the modification. *Id.* § 3604(f)(3)(A)-(B).
103 *Id.* § 3601(1).
104 *Bloch v. Frischholz (Bloch I)*, 533 F.3d 562, 565 (7th Cir. 2008), aff’d in part, rev’d in part, 587 F.3d 771 (7 Cir. 2009) (en banc) (“It would be especially inappropriate to adopt in the name of the Fair Housing Act a principle that lack of accommodation = discrimination, since the FHA itself distinguishes the two. By requiring accommodation of handicap but not race, sex, or religion, the statute’s structure tells us that the FHA uses the word ‘discriminate’ to mean something other than ‘failure to accommodate.’”).
discrimination. By incorrectly granting summary judgment, the Seventh Circuit did not reach this key question.

Case law in the Seventh Circuit is consistent with this narrow interpretation of the language of the Fair Housing Act, but Supreme Court precedent is not. The Seventh Circuit has routinely refused to look beyond the words of the Fair Housing Act to the legislative history to give additional breadth to the statute. In contrast, the Supreme Court has held that the Fair Housing Act should be interpreted broadly and has looked to the legislative history in assessing procedural Fair Housing Act questions. It is reasonable, therefore, to suggest that the congressional intent of the statute may require a broader interpretation of the statute’s language than the words alone may indicate.

The Blochs’ request to display their mezuzah may be construed as a request for an accommodation to a facially neutral rule. Although a right to accommodation is not explicitly conferred in the statute, such a right may be implicit, given the law’s purpose. To begin, the Blochs merely requested the right to display their mezuzah in their doorway. Accommodation of this request would not cost the condominium association any money, present any administrative burden to change the rule, nor make an exception for the display of religiously-mandated objects. In

---

105 See, e.g., Halprin v. Prairie Single Family Homes of Dearborn Park Ass’n, 388 F.3d 327, 330 (7th Cir. 2004) (“Reference to legislative history is criticized when it is used to give a statute a reach that exceeds what its words suggest. Our use here is the opposite; it is to confirm that the words mean what they seem to mean.”). It is a Seventh Circuit trend to advocate for “plain meaning” interpretation of statutes. Judge Easterbrook, who writes for the majority in Bloch I, routinely advocates for narrowing the scope of statutory interpretation with his philosophy of “new textualism.” Easterbrook believes that “it is misleading to speak of legislative intent,” and, as a result, he has advocated for “a limit on the use of legislative history.” James E. Westbrook, A Comparison of the Interpretation of Statutes and Collective Bargaining Agreements: Grasping the Pivot of Tao, 60 Mo. L. Rev. 283, 287-88 (1995). Similarly, Judge Posner has “caution[ed] against judicial reliance upon broad statutory purpose.” Id.


107 The requested accommodation therefore meets the statutory requirements for accommodation as laid out for handicaps. See 42 U.S.C. § 3604(f)(2)(A)-(B).
addition, failure to accommodate this request prevents the Blochs from “enjoy[ing their] dwelling,” further rendering it unusable to them. The words of the Fair Housing Act explicitly state that it is unlawful “[t]o refuse to sell or rent . . . or otherwise make unavailable” a dwelling on the basis of religion. Contrary to the holding in *Bloch II*, the Shoreline Towers Condominium Association’s hallway rule effectively made the condominiums unavailable to observant Jewish individuals who are required by the tenets of their religion to display a *mezuzah* in the doorway of their homes. Although the Blochs did not move out of their condo, as the Seventh Circuit asserts would necessarily need to be shown here for a claim of constructive eviction, the hallway rule’s treatment of observant Jews is discrimination on the basis of religion—the exact conduct the Fair Housing Act prohibits. In fact, despite the finding that the Blochs’ home was not made “unavailable,” the court in *Bloch II* specifically acknowledged that “Section 3604(a) is designed to ensure that no one is denied the right to live where they choose for discriminatory reasons,” a proposition aptly describing the situation facing the Blochs. If a facially neutral rule results in a discriminatory impact on people’s religion, there *should* be a cause of action under the Fair Housing Act, even if no duty exists to accommodate for religious observance.

108 Id. § 3604(f)(3)(B).
109 Id. § 3604(a) (emphasis added). The full text of this clause states that it shall be unlawful, “To refuse to sell or rent after the making of a bona fide offer, or to refuse to negotiate for the sale or rental of, or otherwise make unavailable or deny, a dwelling to any person because of race, color, religion, sex, familial status, or national origin.” Id.
110 Bloch v. Frischholz (*Bloch II*), 587 F.3d 771, 777-78 (7th Cir. 2009); see supra note 62 and accompanying text. The en banc decision asserts that “[w]hether ‘unavailability’ means that a plaintiff must, in every case, vacate the premises to have a § 3604(a) claim is an issue we refrain from reaching.” *Bloch II*, 587 F.3d at 778. Yet despite this statement, that is precisely the onus put on the Blochs, as the court noted “the Blochs never moved out [and] gave no reason why they failed to vacate,” and concluded that “based on these facts, we see no possibility that a reasonable jury could conclude that the defendants’ conduct rendered Shoreline Towers ‘unavailable’ to the Blochs.” *Id.*
111 *Bloch II*, 587 F.3d at 777-78; see supra note 62 and accompanying text.
112 The en banc decision noted that “if the Blochs produced sufficient evidence of discrimination, we conclude that § 3604(b) could support the Blochs’ claim,” yet the court never reached this issue. *Bloch II*, 587 F.3d at 781.
113 *Bloch II*, 587 F.3d at 776 (internal quotation marks omitted) (quoting Southend Neighborhood Improvement Ass’n v. County of St. Clair, 743 F.2d 1207, 1210 (7th Cir. 1984)).
114 See Hack v. President & Fellows of Yale Coll., 237 F.3d 81, 103-04 (2d Cir. 2000) (Moran, J., dissenting) (asserting that although the college may not be compelled
Similarly, the Fair Housing Act makes it unlawful to “coerce, intimidate, threaten, or interfere with any person in the exercise or enjoyment of... any right granted or protected by section 3603, 3604, 3605, or 3606 of [the Act].” Section 3604 makes it unlawful to discriminate against individuals in the sale, rental, and privileges of sale or rental of a dwelling on the basis of religion. Seventh Circuit case law interpreting this clause of the Fair Housing Act has prohibited harassment that amounts to constructive eviction, analogizing such conduct to constructive discharge, which is prohibited under Title VII of the Civil Rights Act. In the case of the Blochs, their inability to display their mezuzah in their doorway was constructive eviction because they were no longer able to live in their condominium and simultaneously adhere to the rules of their religion. At least one other tenant in Shoreline felt similarly, as she moved out and told her neighbors that she had “essentially been evicted from her home.” Because of Shoreline’s persistent removal of the mezuzot from the Blochs’ doorway and prevention of Lynn’s display of a mezuzah while she was mourning the death of her husband, these actions may be classified as harassment and result in a cause of action under the Fair Housing Act, which various courts in the country have allowed.

Finally, although § 3604 does not explicitly address post-acquisition discrimination, hindrance of one’s ability to

by the Fair Housing Act to make religious accommodations, where plaintiff makes a prima facie showing of the unavailability of housing to observant Jewish students, the case should proceed to discovery and be left to the fact finder to determine whether the rule has had a discriminatory effect).


Id. § 3604(a)-(b).

DiCenso v. Cisneros, 96 F.3d 1004, 1008 (7th Cir. 1996). For a more in-depth discussion of the parallels between Title VII constructive discharge claims and the potential for creating a Title VII constructive eviction cause of action, see infra Part III.C.


DiCenso, 96 F.3d at 1008 (noting that several other courts have found harassment to be an actionable form of housing discrimination (citing Beliveau v. Caras, 873 F. Supp. 1393, 1396-97 (C.D. Cal. 1995); People v. Merlino, 694 F. Supp. 1101 (S.D.N.Y. 1988))).

In Bloch II, the Seventh Circuit created a case for a claim of post-acquisition discrimination under § 3604 when it noted that, “[a]s a purely semantic matter the statutory language [of Section 3604(a)] might be stretched far enough to reach a case of constructive eviction.” Bloch v. Frischholz (Bloch II), 587 F.3d 771, 776 (7th Cir. 2009) (internal quotation marks omitted) (quoting Halprin v. Prairie Single Family Homes of Dearborn Park Ass’n, 388 F.3d 327, 329 (7th Cir. 2004)). The court went on to explicitly state that “§ 3604(a) may reach post-acquisition discriminatory
enjoy the rights guaranteed in this section can take place after acquisition of the dwelling, creating a cause of action under § 3617, which prohibits coercion, intimidation, and interference with one’s enjoyment of his dwelling.\textsuperscript{121} While the Seventh Circuit has “routinely reserved” the issue of whether a plaintiff may assert a cause of action under § 3617 even in the absence of violations of §§ 3603, 3604, 3605, or 3606,\textsuperscript{122} at least two courts have found a valid cause of action in these circumstances.\textsuperscript{123} Furthermore, as the United States pointed out in its amicus brief to the Seventh Circuit prior to the en banc re-hearing, the Department of Housing and Urban Development has itself interpreted the Fair Housing Act to apply to “post-acquisition discrimination.”\textsuperscript{124} Affording \textit{Chevron} deference to HUD’s interpretation of the Fair Housing Act forces the conclusion that “post-acquisition discrimination” is protected under the Act.\textsuperscript{125}

Thus, the Seventh Circuit’s consideration of the Blochs’ Fair Housing Act claims in both \textit{Bloch I} and \textit{II} wrongly affirmed the grant of summary judgment because they failed to fully and properly consider the various theories discussed herein that would support a cause of action for discrimination under the Fair Housing Act. A look into the congressional intent of the Fair Housing Act reveals that, in order to achieve the stated goal of the statute, a broader interpretation of the statutory language is necessary.\textsuperscript{126} The discriminatory conduct that makes a dwelling unavailable to the owner or tenant, somewhat like a constructive eviction.” \textit{Bloch II}, 587 F.3d at 776. However, finding no “unavailability” here, the Blochs’ claim under § 3604(a) was dismissed.

\textsuperscript{121} See Halprin, 388 F.3d at 330.

\textsuperscript{122} \textit{Bloch II}, 587 F.3d at 781.

\textsuperscript{123} See United States v. Koch, 352 F. Supp. 2d 970, 978-79 (D. Neb. 2004); Stackhouse v. DeSitter, 620 F. Supp. 208, 210 (N.D. Ill. 1985). Additionally, in \textit{Bloch II}, the Seventh Circuit stated that “a § 3617 claim might stand on its own,” and in applying that idea to the Blochs, noted that a claim under this section could only prevail \textit{after proof} of intentional discrimination, an issue remanded for determination at trial. \textit{Bloch II}, 587 F.3d at 782-83. Once discrimination was established, the Blochs would be able to proceed under § 3617, but that issue has yet to be reached. \textit{Id.}

\textsuperscript{124} Brief for the United States as Amicus Curiae in Support of Plaintiffs-Appellants Urging Reversal and Remand on Fair Housing Act Claims at 32, Bloch v. Frischholz, 533 F.3d 562 (7th Cir. 2008) (No. 06-3376) (citing 24 C.F.R. 100.400(c)(2) and 24 C.F.R. 100.65(b)(4)).

\textsuperscript{125} \textit{Id.} at 33 (citing \textit{Chevron}, USA, Inc. v. Natural Res. Def. Council, 467 U.S. 837 (1984)). \textit{Chevron} sets forth a legal test for determining whether deference should be given to a government agency’s interpretation of their statutory law. \textit{See Chevron}, 467 U.S. 837. Defeference is given to the agency when a statute is ambiguous and the agency’s interpretation is reasonable. \textit{See id.}

\textsuperscript{126} The Supreme Court has recognized the importance of expanding the scope of interpretation of the Civil Right Act of 1964. \textit{See Griffen v. Breechinridge}, 403 U.S.
treatment to which the Blochs were subject is exactly the type of conduct that the Fair Housing Act seeks to prohibit, and by granting summary judgment, the Seventh Circuit effectively undermined the ability of the Fair Housing Act to “provide . . . fair housing throughout the United States.”

B. The Necessity for Religious Accommodation Under the First Amendment

While the Fair Housing Act may not explicitly require religious accommodation, the question of whether the First Amendment requires such an accommodation in the housing context is a separate but equally important issue. As discussed in Part II, the First Amendment is implicated here because housing discrimination occurred on the basis of religion. Analyzing how courts have dealt with religious accommodation in other contexts sheds some light on whether the situation in Bloch was appropriately analyzed under the First Amendment. First Amendment jurisprudence, specifically the religious accommodation requirement of Title VII of the Civil Rights Act, demonstrates that not only would a parallel requirement under the Fair Housing Act be constitutional, but it would also further the Act’s goals. In failing to recognize the necessity of a religious accommodation requirement under the First Amendment, the Seventh Circuit panel in Bloch I rashly and inappropriately granted summary judgment, and only partially nullified the effects of this harsh judgment in Bloch II by remanding the case in part, underscoring the need for action.

The religion line of the First Amendment is generally analyzed as two separate clauses—the Free Exercise Clause and the Establishment Clause. The Free Exercise Clause prohibits the creation of laws that burden the practice of

88, 97-98 (1971) (acknowledging that the Court has broadly interpreted Civil Rights statutes); see also supra note 106 and accompanying text.

127 42 U.S.C. § 3601 (2006). While Bloch II leaves open the possibility for the Blochs to establish claims under § 3604(b) and § 3617 if a finding of intentional discrimination is made on remand, no relief is presently available to the Blochs under the Fair Housing Act without further litigation pending a favoring outcome on remand. Bloch II, 587 F.3d at 781, 783.

128 Id. § 2000e-1 to -17.

129 The Religion Clauses of the First Amendment state, “Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof . . . .” U.S. CONST. amend. I. The first ten words of the Amendment have been deemed the “Establishment Clause” and the remainder the “Free Exercise Clause.” See Everson v. Bd. of Educ., 330 U.S. 1, 32 (1947).
religion, while the Establishment Clause prohibits the
government from promoting or showing favoritism towards one
religion.\textsuperscript{130} When a law places an incidental burden on the
practice of religion, the Free Exercise Clause merely requires
that such a law be neutral and generally applicable in order to
be deemed constitutional.\textsuperscript{131} When this test cannot be met and a
law is not religiously neutral, the government must prove that
the law is necessary to achieve a compelling government
interest and is narrowly-tailored to meet this interest.\textsuperscript{132}
Finally, the government must prove that the achievement of
the interest at stake would be undermined by the creation of a
religious exception to the rule.\textsuperscript{133} Similarly, the Establishment
Clause forbids a law from promoting or disadvantaging any
particular religion.\textsuperscript{134} The Supreme Court has stated that there
are three ways of proving that a law suppresses religion or
religious conduct: (1) the law is facially biased; (2) the law
targets one or more religious groups; or (3) the law prohibits
more conduct than is necessary to achieve the stated
compelling government interest.\textsuperscript{135}

The Shoreline Towers Condominium Association rule
violates the First Amendment’s Free Exercise Clause and may
be deemed unconstitutional under two of the theories that the
Supreme Court has approved for proving that the purpose of a
law is to suppress religious conduct. First, the condominium
rule targets a religious group. While not expressly targeting
Jews, the rule effectively prohibits observant Jews from living
at Shoreline Towers in accordance with the tenets of their
religion. As Justice Wood pointed out in her dissent in \textit{Bloch I},
just as a rule that forbids individuals from wearing
headscarves in the common areas of the condominium would
single out observant Muslim women whose religion requires
them to cover their heads, so too does the prohibition of the
display of religiously mandated objects in the doorways of one’s
home disproportionately affect observant Jews.\textsuperscript{136} Second, the

\textsuperscript{130} U.S. CONST. amend. I.
\textsuperscript{131} See Employment Div. Dep’t of Human Res. v. Smith, 494 U.S. 872, 878-79
\textsuperscript{132} Church of Lukumi Babalu Aye, Inc. v. City of Hialeah, 508 U.S. 520, 530
(1993).
\textsuperscript{133} Id.
\textsuperscript{134} Id. at 532 (citing Bd. of Educ. v. Mergens, 496 U.S. 226, 248 (1990)).
\textsuperscript{135} See id. at 533-38.
\textsuperscript{136} Bloch v. Frischholz (\textit{Bloch I}), 533 F.3d 562, 573 (7th Cir. 2008) (Wood, J.,
dissenting) \textit{aff’d in part, rev’d in part}, 587 F.3d 771 (7th Cir. 2009) (en banc).
hallway rule prohibits more conduct than is necessary to achieve the Association’s stated goal of maintaining an attractive hallway appearance.\footnote{137} This objective could be achieved by less restrictive means. Moreover, it could be achieved even if there was a religious exception to the rule permitting Jews or any other religious group that requires such an accommodation to display mezu\textit{z}ot or the like in their doorways.

In sum, the condominium association rule violates the Free Exercise Clause under the factors enumerated by the Supreme Court. On the other hand, a religious accommodation requirement in the housing context would be constitutionally permissible. In fact, religious accommodation would be in line with the legislative approach to religion in the employment context, a world where religious accommodation is not merely permissible, but is required.\footnote{138} Even if religious accommodation in the housing context is out of line with the congressional intent of the Fair Housing Act, the subsequent sections assert that religious accommodation is necessary, at least in certain situations. These situations exist where, like in the case at bar, the rule in question has a disparate impact on one specific religious group, observant Jews, and is therefore in violation of both the First Amendment and the Fair Housing Act. In such an instance, religious accommodation is necessary regardless of whether the drafters of the Fair Housing Act envisioned the law to operate in this way.

C. Religious Accommodation and Title VII Employment Claims

Title VII of the Civil Rights Act prohibits employment discrimination on the “basis of race, color, religion, sex, and national origin.”\footnote{139} The Supreme Court has held that it is constitutionally permissible for the government to order employers, under this Act, to accommodate employees’ religious

\footnote{137 See Brief of Defendants-Appellees at 9, Bloch v. Frischholz, 533 F.3d 562 (7th Cir. 2008) (No. 06-3376).} \footnote{138 See 42 U.S.C. § 2000e(j) (2006); see also EEOC, RELIGIOUS DISCRIMINATION (2009), http://www.eeoc.gov/types/religion.html [hereinafter RELIGIOUS DISCRIMINATION]. This assertion builds off the previous section’s argument that religious accommodation is in line with the congressional intent of the Fair Housing Act. \textit{See supra} Part III.C.} \footnote{139 42 U.S.C. § 2000e-2(a)(1).}
This section seeks to explain how religious accommodation is handled under the Title VII mandate, and argues that creating a similar requirement under Title VIII, the Fair Housing Act, would function similarly and be constitutionally permissible.

The Supreme Court has held that the Title VII religious accommodation requirement does not violate the Establishment Clause, because it does not promote or advance religion, but rather merely permits religious exercise. The United States Equal Employment Opportunity Commission ("EEOC"), the agency responsible for enforcing Title VII and investigating alleged violations of the statute, created the religious accommodation requirement, which now requires that "[e]mployers must reasonably accommodate employees' sincerely held religious practices unless doing so would impose an undue hardship on the employer." The EEOC has defined a religious accommodation to be "any adjustment to the work environment that will allow the employee to practice his religion." Such adjustments include flexible scheduling, task reassignments, and modification of grooming requirements and agency policies, practices, and procedures, including permitting religious expression. The only exception to this requirement is if such an accommodation would "legitimately" harm the interests of the business.

The adoption of religious accommodation practices in the housing context would be no more violative of the First Amendment than would the parallel requirement in the employment context, which the Supreme Court has long upheld as constitutional. A religious accommodation requirement is necessary to provide homeowners with the same protections that the EEOC requirement of religious accommodation provides to employees. Such an accommodation, as long as it is tailored in a manner similar to Title VII, should not violate the

---

140 See Corp. of Presiding Bishop v. Amos, 483 U.S. 327, 334 (1987) ("This Court has long recognized that the government may (and sometimes must) accommodate religious practices and that it may do so without violating the Establishment Clause.") (internal quotation marks omitted) (quoting Hobbie v. Unemployment Appeals Comm'n, 480 U.S. 136, 144-45 (1987)).

141 See id.

142 RELIGIOUS DISCRIMINATION, supra note 138.

143 Id.

144 Id.

145 Id.

146 See Amos, 483 U.S. at 335.
Establishment Clause. To the contrary, it would further the goal of the Free Exercise Clause of preventing the burdening of religious exercise, because such an accommodation would enable homeowners to be free from the fear that they may be constructively evicted from their homes due to rules that inhibit their ability to practice their religion in their homes.

The Title VII religious accommodation requirement has interpreted the word “religion” to include “all aspects of religious observance and practice, as well as belief.” In order to establish a religious accommodation claim under Title VII, the claim must be assessed under a two-step framework. First, the claimant must establish a prima facie case by proving that: (1) he has a “bona fide religious belief” that conflicts with one or more of his employment duties; (2) he informed the employer of the belief and conflict; and (3) as a result of the conflict, he was subject to adverse employment consequences. If all of these requirements are met, the burden shifts to the defendant to show that accommodation would create an undue burden for the company.

Employers are required to make religious accommodations for both employees and prospective employees, so long as such an accommodation does not result in “undue hardship.” However, the level of undue hardship necessary to overcome the religious accommodation requirement is not clear. The Ninth Circuit found in Garbers v. Postmaster General that forcing an employer to pay extra overtime wages was an undue hardship justifying the employer’s refusal to provide a Baptist employee additional time off to attend ministerial meetings twice a month. That court defined undue hardship as “resulting in more than a de minimis cost to the employer.” Furthermore, the Supreme Court has held that when an employer takes reasonable steps to accommodate an

---

147 U.S. CONST. amend. I.
151 Johnson, 1997 WL 741368, at *1; see also Garbers, 1995 WL 241474, at *1.
152 Trans World Airlines, 432 U.S. at 74.
153 Garbers, 1995 WL 241474, at *1 (holding that having to pay 12% in overtime constituted an undue burden to employer overcoming his duty to accommodate employee’s religious beliefs).
154 Id. (citing Heller v. EBB Auto Co., 8 F.3d 1433, 1440 (9th Cir. 1993)); see also Trans World Airlines, 432 U.S. at 84.
employee’s religious beliefs, the requirement may be met even though no solution was ultimately found.\footnote{155 See Trans World Airlines, 432 U.S. at 78-79 (holding that no violation of Title VII occurred where an airline provided a religious employee with multiple alternatives to working on Saturdays, despite the fact that none were ultimately found to be suitable; the Court found that the accommodation that the employee wanted would have posed an undue burden, and because employee had been given accommodating options, Title VII was not violated).}

In addition, employees may bring Title VII claims alleging “constructive discharge.”\footnote{156 See Garbers, 1995 WL 241474, at *1.} Such a situation occurs when “a reasonable person in the employee’s position would have felt that he was forced to quit because of intolerable and discriminatory working conditions.”\footnote{157 Id. at *2 (quoting Watson v. Nationwide Ins. Co., 823 F.2d 360, 361 (9th Cir. 1987)); see also Johnson v. K-Mart Corp., No. 96-2408, 1997 WL 741368, at *3 (4th Cir. Dec. 2, 1997) (defining constructive discharge as the situation where “an employer deliberately makes the working conditions of the employee intolerable in an effort to induce the employee to quit” (internal citations omitted)).} This type of claim requires a plaintiff to establish that the employer deliberately created intolerable working conditions to compel the employee to quit.\footnote{158 See Johnson, 1997 WL 741368, at *1.} Intent is demonstrated through evidence that the employee’s resignation was a “reasonably foreseeable consequence” of the employer’s failure to act when he becomes aware of the condition.\footnote{159 Id.}

Under Title VII’s framework for religious accommodation, it appears that the Blochs’ claim of discrimination would be strong enough to establish a violation. Under the two-step framework, the Blochs have established a prima facie case of discrimination, and there is no proof that providing religious accommodation would cause any hardship for the Shoreline Towers Condominium Association. First, the Blochs have proven that they have a “bona fide religious belief” that conflicts with the rule in question.\footnote{160 See Garbers, 1995 WL 241474, at *1.} The Jewish religion requires that individuals display 
\textit{mezuzot} in the doorways of their homes at all times.\footnote{161 See supra note 13 and accompanying text.} The hallway rule that Shoreline adopted in the spring of 2004 may be interpreted as including a prohibition on the display of 
\textit{mezuzot} on the exterior doors of the condominiums.\footnote{162 Bloch v. Frischholz (Bloch I), 533 F.3d 562, 567 (7th Cir. 2008) (Wood, J., dissenting) aff’d in part, rev’d in part, 587 F.3d 771 (7th Cir. 2009) (en banc).} This rule necessarily interfered with the Blochs’ practice of their religion, and the Blochs appropriately
contacted Shoreline about the problem, providing them with information explaining the religious significance of a mezuzah.163 Finally, the Blochs suffered adverse consequences as a result of the conflict between the hallway rule and their religious practices.164 The Blochs suffered emotional distress as a direct result of both this conflict and their internal struggle between wanting to live in their home and adhering to the tenets of their religion.165 Furthermore, there was no showing that the condominium association would have suffered any hardship at all,166 let alone an “undue hardship,” by accommodating the religious beliefs of the Blochs and other Jewish unit owners.167 The alleged goal of the hallway rule was to “protect the appearance of the hallways,”168 and there was neither a showing that this goal would be undermined by a religious-accommodation exception to the rule, nor that the condominium association would suffer any hardship as a result.

The Shoreline Towers Condominium Association did not even offer accommodation alternatives with which the Blochs may have been satisfied.169 Instead, they entirely refused to accommodate their religious beliefs.170 One tenant described the

163 Id. As the amicus brief submitted in support of Plaintiff-Appellants notes, many Jews believe they are biblically required to place a mezuzah in the doorway of their home because it is “a sacred piece of iconography meticulously presented in their doorway in accordance with [G-d’s] law.” Brief of Amicus Curiae the Decalogue Society of Lawyers in Support of Plaintiffs-Appellants at 4, Bloch v. Frischholz, 533 F.3d 562 (7th Cir. 2008) (No. 06-3376) (“Because an observant Jew cannot live in a home without a mezuzah, forcibly removing one from an adherent’s doorway is tantamount to eviction.”).
164 Bloch I, 533 F.3d at 567 (Wood, J., dissenting); see also Garbers, 1995 WL 241474, at *2-3.
165 See Bloch I, 533 F.3d at 567 (Wood, J., dissenting); Amended Complaint at ¶ 41-42, Bloch v. Frischholz (N.D. Ill. 2006) (No. 06 C 4472).
166 While arguably Shoreline Towers Condominium Association could assert that mezuzot are not aesthetically pleasing, and thus cause the Association hardship, it seems unlikely that a court could legitimately characterize aesthetic concerns as excessive enough to warrant being deemed “undue hardship.” Trans World Airlines, Inc. v. Hardison, 432 U.S. 63, 84 (1977). This seems especially true in light of seriousness that the Supreme Court has required in order to show that an accommodation poses “undue hardship,” that is, proof that the accommodation would pose “more than a de minimis cost” to the accommodator. See id.
167 Id.
168 Response Brief of the Defendants-Appellees at 9, Bloch v. Frischholz, 533 F.3d 562 (7th Cir. 2008) (No. 06-3376).
169 See supra notes 142-145 and accompanying text.
170 Bloch I, 533 F.3d at 568 (Wood, J., dissenting). In September 2004, Lynne Bloch approached the condominium board with a proposed amendment to the hallway rule that would allow the display of religiously-mandated objects, including mezuzot, on the doorframes, but her proposal was rejected. Id.
effect of the rule on her life by stating that she had “essentially been evicted from her home.”\textsuperscript{171} This allegation of constructive eviction is analogous to claims of constructive discharge, which are actionable under Title VII.\textsuperscript{172} Here, observant Jewish condominium owners were placed in a situation where they felt forced to move out of their homes because of intolerable discriminatory housing practices.\textsuperscript{173} Just as such a claim would be actionable under Title VII, a similar remedy should be available for individuals who are placed in an analogous situation by discriminatory housing practices.

The constitutionality of the religious accommodation requirement of Title VII has been established.\textsuperscript{174} Creating a similar requirement for Title VIII, the Fair Housing Act, would similarly pass the First Amendment’s requirements of neither promoting nor advancing religion, and would foster the free exercise of religion. As discussed, if an analogous framework were in place for assessing violations of the Fair Housing Act as is in place for determining whether violations of Title VII have occurred, the Blochs could potentially have two causes of action. First, it seems that the Blochs would easily be able to make out a prima facie case of discrimination that the defendants would be unable to overcome by proof of undue hardship. In addition, the hallway rule resulted in constructive eviction of Jewish condominium owners because they can no longer display their \textit{mezuzot}, in accordance with the tenets of their religion.\textsuperscript{175} As a result, constructive eviction could potentially be an independent cause of action as well.\textsuperscript{176} Creating a religious accommodation requirement for the Fair Housing Act would properly allow discrimination in the housing market to be overcome by the most fair and efficient means possible. The religious accommodation requirement has fostered non-discrimination in the employment context as

\textsuperscript{171} \textit{Id.}
\textsuperscript{173} \textit{Bloch I}, 533 F.3d at 568 (Wood, J., dissenting); \textit{see also Garbers}, 1995 WL 241474, at *7-8.
\textsuperscript{174} Corp. of Presiding Bishop v. Amos, 483 U.S. 327, 335 (1987).
\textsuperscript{175} \textit{See supra} note 13 and accompanying text.
\textsuperscript{176} The Blochs’ potential constructive eviction claim referenced here would be distinct from claims of constructive eviction under § 3604(a) and § 3604(b) of the Fair Housing Act, discussed in Part III.A, as the claim here would stem from the Religious Accommodation Clause for which this section advocates.
efficiently as possible, and so too could such a requirement in the housing market.

D. The Seventh Circuit’s Failure to Reach Disparate Impact Analysis

The Seventh Circuit further erred in analyzing the situation presented in Bloch by failing to assess the disparate impact that the hallway rule had on observant Jews. While the court in Bloch I completely ignored the potential for a disparate impact claim, in Bloch II the court explained that it believed that the Blochs waived their ability to argue disparate impact, because they did not raise it during the summary judgment phase of the proceedings.\textsuperscript{177} The court acknowledged that the Blochs used the term “disparate impact” in their pleadings and cited the seminal case on disparate impact in their sur-reply, but concluded that it was not enough to entitle them to disparate impact analysis on their claims.\textsuperscript{178} However, if the court had reached the Blochs disparate impact claim and looked at Fair Housing Act and First Amendment jurisprudence, it would have become clear that the Blochs had a legitimate disparate impact claim since the hallway rule had a disparate impact on Jewish residents, substantially burdening their ability to freely practice their religion, without any legitimate reason for doing so. In deciding that the Blochs waived their disparate impact claims, both Seventh Circuit decisions never engaged in enough analysis to recognize that clear issues of material fact existed as to the disparate impact of the hallway rule on observant Jews.\textsuperscript{179}

A Fair Housing Act violation can be established by a showing of disparate impact or disparate treatment.\textsuperscript{180}

\begin{footnotesize}
\begin{enumerate}
\item[177] Bloch v. Frischholz (Bloch II), 587 F.3d 771, 784 (7th Cir. 2009).
\item[178] Id. at 784-85 (“Accordingly, we conclude that the Blochs waived any Arlington Heights disparate impact argument. So the Blochs must proceed on a showing on intentional discrimination.”).
\item[179] By limiting the remanded case to the issue of intentional discrimination and not allowing the Blochs to proceed on the additional claim of disparate impact, one of their potential avenues of relief was made unavailable to them.
\item[180] Huntington Branch, NAACP v. Town of Huntington, 844 F.2d 926, 934-35 (2d Cir. 1988), aff’d, 488 U.S. 15 (1988); see also Hack v. President of Yale Coll., 237 F.3d 81, 96-97 (2d Cir. 2000) (Moran, J. dissenting) (noting that “the existence of disparate impact claim under Title VIII was implicitly confirmed when the Supreme Court announced its decision in Village of Arlington Heights v. Metro. Housing Development Corporation.”) (citing 429 U.S. 252 (1977)). The circuits, however, are not in agreement as to whether mere proof of disparate impact alone is enough to establish a Fair Housing Act violation. While some circuits have found violations of the Act in
\end{enumerate}
\end{footnotesize}
Similarly, a violation of the First Amendment may also be established by showing that a rule is facially discriminatory or has a disparate impact on one or more religious groups.\(^{181}\) Proof of disparate impact requires a totality of circumstances analysis, for example, an investigation into the background of the decision, the sequence of events leading up to the decision, and the factors the decision-maker considered.\(^{182}\) In order to prove discriminatory impact under the Fair Housing Act, there must be a showing that the ultimate effect of the law or rule is disproportionately placed on one group.\(^{183}\) If discriminatory impact is present but discriminatory intent cannot be readily proven, the analysis of whether the rule is discriminatory under the Fair Housing Act becomes more complicated. Although Fair Housing Act claims are dramatically easier to prove when both discriminatory intent and effect can be shown, violations of Title VII, as discussed, are often established based on evidence of discriminatory impact, even without a showing of discriminatory intent.\(^{184}\) Thus, whether the additional requirement of proving discriminatory intent is necessary for establishing a violation of the Fair Housing Act, or for establishing whether the conduct is in violation of the First Amendment, must be assessed.

Under the Fair Housing Act, when discriminatory effect can be established but discriminatory intent cannot, a claim that a violation has occurred is not necessarily void. Instead, it just becomes much more difficult to prove.\(^{185}\) The Act requires that in order for a violation to be found in such a case, the

---


\(^{182}\) See Vill. of Arlington Heights, 429 U.S. at 267-68.

\(^{183}\) United States v. City of Black Jack, Mo., 508 F.2d 1179, 1186 (8th Cir. 1974) (holding that discriminatory effect was established where an ordinance led to foreclosure for 85 percent of the African Americans living in the area).

\(^{184}\) United States v. City of Chi., 549 F.2d 415, 435 (7th Cir. 1977) (citing Wash. v. Davis, 426 U.S. 229, 238-39 (1976)).

\(^{185}\) See City of Black Jack, Mo., 508 F.2d at 1186; see also supra note 180. But see Hack, 237 F.3d at 96-97 (Moran, J. dissenting) (citing Resident Advisory Bd. v. Rizzo, 564 F.2d 126, 128 (3d Cir. 1977) and noting that the Third Circuit has held that proof of discriminatory effect alone is sufficient to establish a prima facie case of disparate impact under the Fair Housing Act).
claimant must establish that the action was taken “because of religion.” The broad view of this requirement is that an action can be shown to have been taken on the basis of religion wherever claimant can prove that the “natural and foreseeable consequence” of the action is discrimination, regardless of intent. The narrow view, however, is that in order to meet this requirement, intent needs to actually be proved. Even after this requirement is met, some courts proceed to consider four factors to fully determine whether a violation of § 3604(a), the section prohibiting discrimination in the sale and rental of dwellings, has occurred. The factors to be assessed are: (1) the strength of the showing of the discriminatory effect; (2) whether there is any showing of discriminatory intent at all; (3) the defendant’s interest in taking the action; and (4) whether the plaintiff is looking to get the court to make an affirmative holding that the defendant must provide housing for minorities or whether he seeks merely to prevent the defendant from interfering with the right of individual property owners.

First Amendment analysis of whether facially neutral laws are discriminatory on the basis of religion is approached differently. Facially neutral laws are assessed under a rationality standard: does the rule infringe on the free exercise of religion? Even if a facially neutral law of general applicability does infringe on religious practice, the Supreme Court has only applied strict scrutiny where other constitutional protections were at issue as well, or where it could be shown that the rule was not actually “neutral,” because its object was “to infringe upon or restrict practices.”

---

187 See id. at 1288.
188 See id.
189 See id. at 1290.
190 Id.
192 See id. at 881-82 (citing numerous cases including West Virginia Bd. of Educ. v. Barnette, 319 U.S. 624 (1943) (deeming a statute compelling students to salute the flag to be unconstitutional when challenged by religious objectors); Wisconsin v. Yoder, 406 U.S. 205, 206 (1972) (invalidating compulsory schooling laws as applied to Amish Parents who objected on religious grounds after the Court upheld the right of parents to direct the education of their children in Pierce v. Soc’y of Sisters, 268 U.S. 510 (1925)).
Congress, however, did not approve of the rational basis test, and voiced its disapproval by passing the Religious Freedom Restoration Act ("RFRA").

RFRA prevented the government from "substantially burdening" religious practice even where the burden resulted from a generally applicable rule. The exception to this was that if the government could demonstrate that the burden was in furtherance of a "compelling government interest," and that the restriction was the least restrictive means of achieving the interest, then the rule could be upheld. One of the stated goals of RFRA was to provide a cause of action for people whose religious exercise was unnecessarily burdened by the government. The Supreme Court, however, struck down RFRA in City of Boerne v. Flores, holding that the law created a separation of powers issue because it exceeded Congress' remedial powers.

In a second attempt to remedy what Congress saw as an injustice in denying individuals the rights guaranteed to them by the First Amendment, Congress passed the Religious Land Use and Institutionalized Persons Act ("RLUIPA"). This Act was also constitutionally challenged, but the Supreme Court upheld the statute, exemplifying the fact that religious accommodation statutes do not automatically violate the Establishment Clause. RLUIPA is a narrower version of RFRA that prohibits the government from imposing "substantial" burdens on the religious exercise of any institutionalized individual unless there is a compelling government interest and the restriction is the least restrictive means of achieving the interest. In the face of allegations that RLUIPA

194 42 U.S.C. §§ 2000bb-1 to -4 (1993) (this statute's application to local and state governments was later declared unconstitutional in City of Boerne v. Flores, 521 U.S. 507 (1997)).
195 Id. § 2000bb.
196 See City of Boerne, 521 U.S. at 515-16 (discussing the intent of RFRA before deeming it unconstitutional).
199 City of Boerne, 521 U.S. at 536 (noting that "RFRA contradicts vital principles necessary to maintain separation of powers and the federal balance"); see also Cutter v. Wilkinson, 544 U.S. 709, 715 (2005) (explaining that "this Court invalidated RFRA as applied to States and their subdivisions, holding that the Act exceeded Congress' remedial powers under the Fourteenth Amendment").
201 See Cutter, 544 U.S. at 719-26 (holding that the protections granted by RLUIPA did not violate the Establishment Clause).
202 See 42 U.S.C. § 2000cc-1(a)(1)-(2); see also Cutter, 544 U.S. at 712.
effectively advanced religion, in conflict with the Establishment Clause, the Supreme Court upheld the statute. In *Cutter v. Wilkinson*, the Supreme Court found RLUIPA to be constitutional because there was “no cause to believe RLUIPA would not be applied in an appropriately balanced way” given the “compelling governmental interest” exception, and the fact that it “does not differentiate among bona fide faiths.” The Court recognized RLUIPA to be a continuation of “congressional efforts to accord religious exercise heightened protection from government-imposed burdens, consistent with this Court’s precedents.” Thus, the Supreme Court adeptly observed Congress’ discontent with the limited protections that the rationality standard advanced in *Employment Division v. Smith* provided to religious individuals in the face of governmental action stymieing individuals’ ability to practice freely.

In *Bloch I*, the Seventh Circuit found the hallway rule to be neutral and generally applicable with respect to religion. Viewing the rule as facially neutral and seeing no requirement for religious accommodation under the Fair Housing Act, the court took their analysis no further. By stopping their analysis at this early stage, the opinion failed to determine whether the hallway rule had a disparate impact on individuals of particular religious groups. The *Bloch II* court recognized that while the hallway rule may have been neutral when first adopted, the crux of the Blochs’ claim was that the reinterpretation of the rule in 2004 was not neutral. Yet the court never engaged in disparate impact analysis to address this, and remanded the case only for determination of intentional discrimination. But, violations of both the Fair Housing Act and the First Amendment can be made by proving that a disparate impact results from application of a facially

---

204 *Cutter*, 544 U.S. at 722.
205 Id. at 723.
206 Id. at 714.
209 *Bloch v. Frischholz (Bloch I)*, 533 F.3d 562, 564 (7th Cir. 2008), *aff’d in part, rev’d in part*, 587 F.3d 771 (7th Cir. 2009) (en banc).
210 Id. at 564-65.
211 *Bloch v. Frischholz (Bloch II)*, 587 F.3d 771, 783 (7th Cir. 2009).
neutral rule. Advanced analysis of the implications of the hallway rule sheds light on the fact that perhaps disparate impact analysis would have led to a different outcome.

Under Fair Housing disparate impact analysis, even if we begin with the assumption that discriminatory intent is not outwardly evident, when we look to the factors for determining whether a violation has occurred, the answer appears to be overwhelming. First, there is a strong showing of discriminatory effect here as the hallway rule results in the constructive eviction of Jewish residents since they can no longer display their mezuzot. This effect is significant as it inhibits observant Jews from living in the condominiums they have purchased, and effectively creates a situation where the Shoreline Towers Condominiums are able to exclude Jewish residents. Second, discriminatory intent is not absent here. Rather, the record is replete with evidence of discriminatory intent. The fact that the hallway rule was in place for three years before the association decided to change its applicability to include the display of mezuzot was clearly a targeted action. The Association began to remove and confiscate mezuzot without giving notice to the residents of the new interpretation of the hallway rule, and continued this practice, eventually culminating in a threat to the Blochs that affixation of a mezuzah would result in a monetary penalty. In addition, the fact that the display of mezuzot does nothing to inhibit the stated goal of the hallway rule, to protect the hallways' appearance, is further evidence of the rule's discriminatory intent.

---

214 See id.; see also supra note 13 and accompanying text.
216 Bloch v. Frischholz (Bloch I), 533 F.3d 562, 567 (Wood, J., dissenting), aff’d in part, rev’d in part, 587 F.3d 771 (7th Cir. 2009) (en banc); Bloch II, 587 F.3d at 773 (“Though Frischholz knew as early as 2001 that removing mezuzot would be a problem for Lynne Bloch, he made no effort to stop the staff from repeatedly tearing them down. Instead, he accused Lynne of being a racist, called her a liar, encouraged other tenants to vote against her reelection to the Association’s Board of Managers, and told her that if she didn’t like the way the rules were enforced, she should ‘get out.’”).
217 Bloch I, 533 F.3d at 567.
218 Response Brief of Defendants-Appellees at 9, Bloch v. Frischholz, 533 F.3d 562 (7th Cir. 2008) (No. 06-3376).
To add insult to injury, the condominium association removed the Blochs’ mezuzah after they requested special permission to display one only for a limited time to conform to the Jewish mourning rituals after the death of Marvin Bloch. The condominium association’s action demonstrates that another motivation, other than the hallway’s appearance, was at issue. Finally, the fact that the condominium board rejected a proposal to create a religious accommodation amendment to the hallway rule, when such an amendment would have no impact on the achievement of the rule’s purpose, seems to indicate that they may have had an ulterior motive. This rejection is additional evidence of the association’s discriminatory intent in interpreting the hallway rule to prohibit even the display of religiously-mandated objects.

Looking to the final two factors, it seems clear that not only did Shoreline lack an interest in taking this action, but also that all the Blochs were seeking to achieve with their lawsuit was to prevent interference with their individual property rights. Shoreline cannot legitimately say that the amendment proposed by the Blochs was objectionable on any sound basis. If the hallway rule’s purpose really is simply to protect the appearance of the hallways, then the display of small religious objects such as a mezuzah does not contradict or undermine this purpose in any way. This is a question of fact that remained to be brought before a jury. Thus, the defendants had no interest in taking this action. Finally, the Blochs did not request that the Court require Shoreline to take an affirmative action to provide housing for religious minorities. Rather, they merely sought to protect their own individual property rights, which they were prevented from

219 Bloch I, 533 F.3d at 567 (Wood, J., dissenting) ("Perhaps the worst episode, and one that gives rise to a strong inference of anti-Semitic animus occurred while the Blochs were mourning the death of Dr. Marvin Bloch, Lynne's husband and Helen and Nathan's father.").
220 Id. at 568.
222 Brief for the United States as Amicus Curiae in Support of Plaintiffs-Appellants Urging Reversal and Remand on Fair Housing Act Claims at 37, Bloch v. Frischholz, 533 F.3d 562 (7th Cir. 2008) (No. 06-3376) (explaining that "were the jury to find that the Association’s actions were motivated in part by plaintiffs’ race or religion, it does not matter that the Association acted under the aegis of neutral Hallway Rule 1").
223 See id.
exercising.\textsuperscript{224} Given the analysis of these factors, a violation of the Fair Housing Act has likely occurred, even without a reading of the statute that requires religious accommodation, because of the disparate impact of the hallway rule on observant Jewish residents and the fact that HUD has interpreted the Act to apply to post-acquisition discrimination.\textsuperscript{225}

Under the First Amendment, while the seemingly facially neutral and generally applicable hallway rule at issue in \textit{Bloch} might initially be seen as a situation warranting only rational-basis review,\textsuperscript{226} this superficial analysis oversimplifies the complexity of the issues raised in the case. As the Supreme Court’s decision in \textit{Church of Lukumi Babalu Aye, Inc. v. City of Hialeah}\textsuperscript{227} makes clear, strict scrutiny must be applied where the purpose of a law is to impede the practice of religion.\textsuperscript{228} In his concurring opinion, Justice Scalia expounds on this, noting that strict scrutiny is required where “laws which, though neutral in their terms, through their design, construction, or enforcement, target the practices of a particular religion for discriminatory treatment.”\textsuperscript{229} That is precisely the type of rule at issue here.\textsuperscript{230} The hallway rule, while neutral on its face, has been enforced only to the detriment of observant Jews, who are no longer free to display their \textit{mezuzot}. Where the application of a facially neutral law reveals religious animus and results in disparate impact on only one religious group, the law can no longer be deemed “neutral.”\textsuperscript{231} Thus, the Condominium Association’s interpretation of the hallway rule to include a prohibition on the display of \textit{mezuzot}, combined with their

\begin{footnotesize}
\begin{enumerate}
\item[224] See id.
\item[225] See supra note 50 and accompanying text.
\item[228] \textit{Church of Lukumi Babalu}, 508 U.S. at 533 (“Although a law targeting religious beliefs as such is never permissible, if the object of a law is to infringe upon or restrict practices because of their religious motivation, the law is not neutral, and it is invalid unless it is justified by a compelling interest and is narrowly tailored to advance that interest.” (internal citations omitted)).
\item[229] Id. at 557 (Scalia, J., concurring).
\item[230] In fact, the \textit{Bloch II} court recognized this as they cited to this case in their discussion of intentional discrimination noting, “The First Amendment forbids subtle departures from neutrality’ and ‘covert suppression of particular religious beliefs.”’ \textit{Bloch v. Frischholz (Bloch II)}, 587 F.3d 771, 785 (7th Cir. 2009) (en banc) (citing \textit{Church of Lukumi Babalu}, 508 U.S. at 534). This statement, however, has heavy implications for assessing just how “neutral and generally applicable” the hallway rule actually is.
\item[231] \textit{Church of Lukumi}, 508 U.S. at 533 (majority opinion).
\end{enumerate}
\end{footnotesize}
continual removal of the Blochs’ mezuzah despite specific requests to display it, demonstrate a potential religious motivation behind the rule.232 The situation, therefore, warrants strict scrutiny. When subject to that exacting standard, the hallway rule would certainly fail to overcome the requirement that the rule be justified by a compelling interest and narrowly tailored to meet that interest.233 If the Association’s goal was truly to rid the hallways of clutter, there are certainly less restrictive ways of obtaining this result, including allowing for a religious exception to the rule. It is likely then, that had the Seventh Circuit looked beyond the face of the rule, it would have realized that the situation at bar was analogous to that at issue in Church of Lukumi Babalu Aye, Inc. and therefore warranted a greater level of scrutiny and deeper analysis than it was given.

In acting on its concern for the leniency with which the constitutionality of laws that infringe on religion are assessed, Congress passed legislation that limited the ability of the government to infringe on individuals’ exercise of religion, which was scrutinized by the Supreme Court and upheld.234 In Cutter, the Supreme Court observed that RLUIPA does not “elevate accommodation of religious observances over an institution’s need to maintain order and safety . . . accommodation must be measured so that it does not override other significant interests.”235 It seems logical, therefore, that there could be and should be a statutory requirement for religious accommodation in housing as long as it does not undermine a significant interest, for example, with a compelling interest exception and without giving enhanced privileges to one religion over another.236 Such a requirement would not violate the Establishment Clause and would achieve the objective of the Free Exercise Clause, allowing individuals to freely practice their religions.237

232 See supra notes 214-219 and accompanying text.
233 Church of Lukumi Babalu, 508 U.S. at 533.
234 See discussion of RLUIPA supra notes 198-206 and accompanying text.
236 See id. at 723-33.
237 U.S. CONST. amend. I.
IV. THE ROAD FORWARD: THE NEED FOR IMMEDIATE LEGISLATIVE ACTION

The Seventh Circuit’s cursory ruling in *Bloch I* essentially opened the door for private housing associations to arbitrarily create neutral and generally applicable\(^\text{238}\) rules that only affect one group of people and, in doing so, effectively keep protected classes out of their units. This result is in direct conflict to the stated objectives of the Fair Housing Act, and if taken to the extreme, could cause the nation to revert to segregated housing systems like those that were predominant at the time the statute was enacted.\(^\text{239}\) While the circuit partially remedied these potentially devastating effects by remanding the case for a determination on the issue of intentional discrimination, the initial ruling is one that courts in other circuits could easily follow and the courts’ denial of the Blochs’ § 3604(a) claims still allows great leeway for housing owners to subvert the spirit of the Fair Housing Act. The decision, however, is not without remedy. In light of the harsh consequences that this decision could have on the future of fair housing in the United States, the state and federal legislatures are in the best position to prevent religious discrimination in housing of the type that plagued the Blochs. In fact, some states have already begun to take action, which should encourage other states to become involved in remedying the potentially damaging consequences. In the face of judicial decisions that have negative policy implications, the other branches of government are the only bodies in a position to curtail the effects of such decisions on the general welfare of society.

In the past, when Congress has disapproved of judicial interpretations of the Fair Housing Act and the First Amendment, it has taken steps to overcome these decisions and promote the ideals they originally sought to achieve.\(^\text{240}\) The

---

\(^{238}\) *Bloch v. Frischholz* (*Bloch I*), 533 F.3d 562, 564 (7th Cir. 2008), *aff’d in part, rev’d in part*, 587 F.3d 771 (7th Cir. 2009) (en banc) (citing *Employment Div. v. Smith*, 494 U.S. 872, 872 (1990)).

\(^{239}\) See *supra* note 69 and accompanying text.

Seventh Circuit’s holdings in *Bloch I* and *Bloch II* have extreme policy implications for the future of private housing associations and their ability to discriminate against particular classes, namely religious groups. Since statutory law can effectively prevent these harsh effects from becoming reality, the legislature is in the best position to prevent the discrimination in the housing market that could result from this decision.\(^{241}\) Although the decision may be merely persuasive in jurisdictions outside of the Seventh Circuit, there is no reason why other jurisdictions will not follow the Seventh Circuit’s lead in creating an incredibly low burden for determining whether a Fair Housing Act or First Amendment violation has occurred.\(^{242}\)

The legislatures in some jurisdictions have already taken action to prevent the effects of this decision from creating a situation contrary to that envisioned by the drafters of the Fair Housing Act. While the *Bloch* case was pending in the district court, Chicago enacted an ordinance that prohibits residential building owners from restricting the placement of religious objects in the doorways of homes unless the individual creating the restriction can prove that such a restriction is necessary to avoid property damage or undue hardship to other unit owners.\(^{243}\) Similarly, two years later, Illinois adopted a law requiring condominium associations to reasonably accommodate “religious practices, including the attachment of religiously mandated objects to the front-door area of a condominium unit.”\(^{244}\) Additionally, despite the fact that the *Bloch I* decision was not binding precedent in New York, legislators in the state have proposed a bill that would effectively overturn the *Bloch I* decision and affirmatively

\(^{241}\) See 42 U.S.C. § 3601; see also Otero v. N.Y. City Hous. Auth., 484 F.2d 1122, 1134 (2d Cir. 1973).

\(^{242}\) In fact, just months after the Seventh Circuit rendered its decision, *Bloch* was cited by a district court in the Southern District of New York for the proposition that, “creating a rule that equates failure to accommodate with discrimination would be particularly inappropriate in the context of the FHA, which explicitly provides for accommodation of handicap, but not race, sex or religion.” Ungar v. N.Y. City Hous. Auth., No. 06-Civ.-1968, 2009 WL 125236, at *17 (S.D.N.Y. Jan. 14, 2009).

\(^{243}\) See *Bloch I*, 533 F.3d at 564; CHICAGO, ILL., MUN. CODE 05-8-030(H) (2005).

\(^{244}\) See *Bloch I*, 533 F.3d at 564; 765 ILL. COMP. STAT. 605/18.4(h) (2007).

Representative Jerrold Nadler, the voice behind the proposed New York legislation, also brought the Freedom of Religious Expression in the Home Act of 2008 (FREHA) before the House of Representatives in September 2008,\footnote{Freedom of Religious Expression in the Home Act of 2009. H.R. 6932, 110th Cong. (2007), available at http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_bills&docid=f:h6932ih.txt.pdf.} continuing his crusade to voice disapproval with the Seventh Circuit's precedent-setting decision in *Bloch I*. The bill went before the House, with twenty-four co-sponsors, but was never enacted into law.\footnote{Id.} If enacted, FREHA would have effectively addressed the Blochs' problem by “amend[ing] the Fair Housing Act to prevent discrimination relating to the display of religious symbols.”\footnote{Id.} This proposed amendment to the Fair Housing Act would allow individuals to display religious symbols in their homes, unless prohibition on such displays was “reasonable and necessary to prevent significant damage to property, physical harm to persons, a public nuisance or similar undue hardship.”\footnote{Id.} On September 17, 2008, FREHA was referred to the House Judiciary Committee, and currently remains a potential source of relief for the near future.\footnote{Freedom of Religious Expression in the Home Act of 2009. H.R. 6932, 110th Cong. (2007), available at http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_bills&docid=f:h6932ih.txt.pdf.}

**CONCLUSION**

*Bloch v. Frischholz* brought to light the immense impact that hasty decisions can have by setting important precedents...
with harsh realities. Although the Seventh Circuit attempted to ameliorate the potentially devastating consequences that the *Bloch I* decision could have had, the possibility of adverse implications has not been dispelled. If the Blochs’ intentional discrimination claim does not prevail on remand, then the initial threat of condoning discriminatory housing practices, which the *Bloch I* decision presented, still exists. Additionally, regardless of the outcome on remand, the potential for this result remains because the *Bloch* decision will be precedent in just one circuit. Such a reality comes in spite of the existence of the Fair Housing Act, which has promoted non-discriminatory housing practices throughout the United States since its enactment forty years ago. In the past, congressional action has been, and continues to be, the best means of overcoming judicial decisions that are contrary to the ideals of the nation. Congress should explicitly create a narrowly tailored religious accommodation clause in the Fair Housing Act. The legislative intent of the Fair Housing Act seems to speak to the fact that religious accommodation is necessary to achieve the ideals of the Act. Further reference to the treatment of religious accommodation under the First Amendment in parallel contexts gives an indication of how such a requirement could be developed in a way that would be both constitutionally permissible and potentially further the goals of the Fair Housing Act.

A narrowly tailored religious accommodation requirement in the Fair Housing Act would not violate the First Amendment. Rather, such a requirement would walk the fine line between noninterference with religion, as required by the Free Exercise Clause, and a separation of church and state, as required by the Establishment Clause. The Supreme Court itself has recognized that there is space for legislative action that is not “compelled by the Free Exercise Clause nor prohibited by the Establishment Clause.”

This Note has argued that legislative action creating an accommodation requirement under the Fair Housing Act falls into this “space” and would be the most effective means of overcoming and preventing the potential consequences of the

---

252 Id.
Seventh Circuit’s decisions in *Bloch v. Frischholz*. The complementary goals of the Fair Housing Act and the First Amendment have the ability to “ensure that all persons live comfortably together in our pluralistic society and that all persons have access to safe, decent, sanitary housing where they can exercise their right to worship or not worship as they choose.” With this exceptional ability right at our fingertips, failing to achieve this ideal would wreak havoc on the progress made since the enactment of the Fair Housing Act. The law affords all citizens the ability to create change and live in a nation where they are free from religious persecution and discrimination. Upholding a decision that effectively subverts this notion is unjust and divergent from the ideals upon which this nation was founded.

_Chloe M. Jones^

---

253 *Bloch v. Frischholz (Bloch I)*, 533 F.3d 562 (7th Cir. 2008), _aff’d in part, rev’d in part_, 587 F.3d 771 (7th Cir. 2009) (en banc); *Bloch v. Frischholz (Bloch II)*, 587 F.3d 771 (7th Cir. 2009); _see also Cutter*, 544 U.S. at 719.

254 Seng, _supra_ note 1, at 38.

^ J.D., Brooklyn Law School, 2010; B.A., Economics, Haverford College, 2007. The author would like to thank her parents and sister for their endless love, support, and grammatical guidance throughout this process. The author is also grateful to Professor Joel Gora for his insightful suggestions in the formation of this note topic and to the entire Brooklyn Law Review staff for their diligence editing and publishing this Note. In particular, the author would like to extend extreme gratitude to Andrei Takhteyev, Joseph Roy, Celine Chan, Bill Vandivort, and Matthew Handler for their editorial prowess and tremendous assistance, without which this Note would not have been possible.
The Short-Changing of Investors

WHY A SHORT SALE PRICE TEST RULE IS NECESSARY IN TODAY'S MARKETS

I. INTRODUCTION

The credit crisis that began in the United States in 2007 gripped the world financial system by September 2008, eventually leading to a global recession into 2009, and increased scrutiny of the governmental regulation of financial markets. In the United States, a particular focus was placed on the short selling of equity securities, especially the stocks of financial sector companies most affected by the credit crisis.


2 See Kara Scannell, The Financial Crisis: SEC Issues Short Selling Rules in Bid to Stop Manipulation, WALL ST. J., Sept. 18, 2008, at A6. The credit crisis began in the late 1990’s as a result of a global increase in the availability in credit, spurred by investment in real estate. David Leonhardt, Can’t Grasp Credit Crisis? Join the Club, N.Y. TIMES, Mar. 19, 2008, at A1. The United States housing market caught fire and many of these mortgages, and later other types of debt, were packaged into investment securities called collateralized debt obligations (“CDOs”) and sold to investors globally. Id. Some of these CDOs consisted of subprime mortgages that had higher interest rates because the loans were made to less creditworthy individuals. Id. Many investors, including banks and investment firms, who purchased these securities used high levels of leverage to invest in these CDOs. Id. When the U.S. housing market started to decline in early 2007, some of the subprime borrowers began to default on mortgage payments, which in turn meant that the CDOs purchased by investors also turned into bad investments because investors stopped receiving payments on the mortgages associated with CDOs. Id. The decrease in the value of the assets forced investors to write down the value of these assets, further increasing the investor’s leverage. As the financial system began to de-leverage, banks and other lenders were less willing to
Short selling is the “sale of a security which the seller does not own or any sale which is consummated by the delivery of a security borrowed by, or for the account of, the seller.” A short seller will then return the borrowed shares “by purchasing equivalent securities on the open market.” A short seller expects to profit by purchasing the replacement shares at a price lower than the price at which he sold the borrowed shares. The financial news media speculated as to whether short sellers participated in market manipulation that drove down stock prices and possibly accelerated the demise of several large and established financial corporations including American International Group and Lehman Brothers.

Following these accusations of manipulation, the Securities and Exchange Commission (“SEC”) and the New York State Attorney General announced in September 2008 that they would begin an inquiry into possible “short selling abuses.” At the height of the financial crisis, and in reaction to cascading stock prices as well as continued public scrutiny and speculation concerning market manipulation by short sellers, the SEC took unprecedented action and issued an emergency temporary ban on any short sales of the securities of 799 financial companies.

extend credit to borrowers; investors providing capital began pulling money out because of a fear of being exposed to risky investments like CDOs. Id. The hesitance of banks and other lenders to extend credit to borrowers, even worthy ones, began to affect other areas of the economy. Id. The credit crisis and subsequent economic fallout are much more complicated than portrayed here. This simple summary is merely meant to illustrate why the stocks of financial institutions were particularly susceptible to short selling – the outlook for financial stocks in September 2008 was grim.

5 See id.
7 Bajaj & Glater, supra note 6. New York State Attorney General Andrew Cuomo compared short sellers to “looters after a hurricane.” Id. (internal quotation marks omitted).
This extreme action banning short sales of certain securities was notable because it was a striking divergence from the actions the SEC took just a year earlier, when, in July 2007, it repealed the long-standing Rule 10a-1 under the Securities Exchange Act of 1934. Enacted in 1938 and commonly known as the “uptick rule,” Rule 10a-1 was intended “to restrict short selling in a declining market.” The rule prohibited short selling a stock at a price less than the price of the “immediately preceding” sale of that stock. The SEC rescinded Rule 10a-1 in 2007 to “modernize and simplify short sale regulation,” judging the rule “no longer . . . effective or necessary.” The SEC’s repeal of the uptick rule was, therefore, a complete repudiation of its longstanding judgment that unrestricted short selling could be dangerous in a falling stock market.

The stated mission of the modern-day SEC is to “protect investors, maintain fair, orderly, and efficient markets, and...
facilitate capital formation.” The Commission was formed in 1934 after the devastating effects of the stock market crash of 1929, and almost immediately, it examined the role of short selling in securities markets and recommended regulation of the practice. In the aftermath of the financial crisis of 2008, the SEC again reconsidered its regulation of short selling and whether the uptick rule or some other short sale price test is necessary to protect investors and preserve orderly financial markets. In April 2009, the SEC solicited public comments to reconsider whether the uptick rule was necessary and if such a rule “would help promote market stability and restore investor confidence.” The SEC also sought comment on whether a modified version of the uptick rule would be more appropriate. The SEC announced an additional public comment period to address “alternative approaches” to the uptick rule in August 2009. Finally in February 2010, the SEC implemented Rule 201, a modified version of the original uptick rule. This version of the rule, which became effective in May 2010, will only be triggered “if the price of an individual security declines intra-day by [ten percent] or more from the prior day’s closing price for that security.” Once this ten percent decline occurs,

16 SEC, FIRST ANN. REP. OF THE SEC 16 (1935), http://www.sec.gov/about/annual_report/1935.pdf. [hereinafter FIRST ANNUAL REPORT OF THE SEC] (“A detailed analysis of the subject of short selling was made for the purpose of determining the extent to which such selling is economically justified and the extent to which it should be curbed.”). At this time, the Commission also recommended to the exchanges that it should implement a version of the uptick rule, believing such a rule would “preserve those features of short selling which are in the public interest.” Id. (emphasis added).
17 When SEC Chairman Mary L. Schapiro was nominated by President Barack Obama in January 2009, she cited re-examination of the uptick rule by the SEC as part of her agenda during her confirmation hearing with the Senate Banking Committee. Stephen Labaton, S.E.C. Nominee Offers Plan for Tighter Regulation, N.Y. TIMES, Jan. 16, 2009, at B3.
19 Id.
22 Id. at 11,234.
short sellers may not sell a security “at or below the current national best bid” for the “remainder of the day and the following day.”

The adoption of this alternative, compromise rule reflects both the SEC’s acknowledgment of the need for some regulation of short selling, and the tentativeness with which the Commission has proceeded in this area, having changed its position three times in just three years. Without attempting to assess the pros and cons of the new rule, this Note argues more generally that regulation of short selling—a feature of our regulatory regime for over seventy years—is necessary to ensure orderly markets and investor confidence. While the adoption of Rule 201 is undoubtedly a positive step, a look at the convoluted history of the rise and fall of the uptick rule reveals a deeper concern over existing justifications for the rule and what degree of short selling regulation is sufficient and appropriate.

This Note will discuss the practice and history of short selling and the uptick rule, including the reasons why the SEC repealed the rule in 2007, whether its removal was in line with the SEC’s mission, and why the rule is necessary to maintain orderly markets. Part II describes the mechanics of short selling equity stocks, the reasons for short selling and why some have a negative view of the practice, and surveys the history surrounding the implementation of the uptick rule in 1938. Part III examines the reasons why the SEC felt the uptick rule was no longer necessary for effective market regulation and the environment in which the SEC made this decision. Part IV discusses alternative statutory provisions the SEC could use to regulate short selling in the absence of the uptick rule. Finally, Part V argues why the reinstatement of the uptick rule or other price test on short selling as a backstop method of protecting investors is necessary because of the difficulty of proving fraudulent or collusive price manipulation through short selling. It also examines the different versions of price tests evaluated by SEC and the new alternative uptick rule that took effect in May 2010. Finally, this Part posits that the uptick rule can be used as a tool to preserve investor confidence and maintain order in troubled, declining markets, preventing both panic and any repeat of the type of radical

measures taken by the SEC in September 2008, when it banned all short sales of financial stocks.

II. THE MECHANICS AND HISTORY OF SHORT SELLING AND THE UPTICK RULE

Concerns and skepticism about the effects of short selling of securities have led to regulation of the practice around the world for hundreds of years. In 1922, J. Edward Meeker, an economist for the New York Stock Exchange, observed that although “prejudice against short selling of securities is not new,” the practice of short selling has “stood that hardest of all tests—the test of time.”

A. The “Mysterious” Practice of Short Selling: How It Works

A short sale is, in simplest terms, a bet that the price of a particular stock will fall. Investors initiate a short sale by borrowing an amount of stock and selling it in anticipation that they will be able to repurchase the same stock later, but at a lower price. The borrower/investor (or his broker) later repurchases the stock at the lower price, returns it to the lender, and profits from the difference between the sale and repurchase price. The borrower also pays the lender a fee for

---

24 J. EDWARD MEEKER, THE WORK OF THE STOCK EXCHANGE 96-97 (1922), available at http://books.google.com/books?id=KDBIAAAAMAAJ&printsec=frontcover&q=The+work+of+the+stock+exchange. The practice of short selling was prohibited in England in 1733, only to be reinstated in 1858. Id. France, the state of New York, and Germany have also passed and later repealed legislation forbidding the practice. Id. After Germany instituted a rule banning short sales on the Berlin Stock Exchange in 1896, at least one economist blamed the rule for an outflow of capital to other international markets leading to depressed markets in Berlin. Id. at 97.

25 Id. at 96. Because short selling is a somewhat sophisticated investment strategy that is not well known to the general public, short selling has also been described as an activity “cloaked in secrecy.” Gary Weiss, The Long and Short of Short-Selling, BUS. WK., June 10, 1991, at 106.

26 Id. at 97.

27 Jonathan R. Macey et al., Restrictions on Short Sales: An Analysis of the Uptick Rule and Its Role in the View of the October 1987 Stock Market Crash, 74 CORNELL L. REV. 799, 799-800 (1989). While different types of assets can be sold short, this Note will focus on the short sale of equity securities.

28 Id. at 799; see also David Chung, Making Sense Out of Market Sentiment Indicators, INVESTOR'S BUS. DAILY, June 16, 1999, at A3; Weiss, supra note 26.

29 Macey et al., supra note 27, at 799-800. In the wake of the financial crisis of 2008, there also was much discussion by business and media commentators about the impact on financial markets of so-called naked short sales. A naked short sale operates similarly to the short sale described here; the major difference is that in a
the use of the stock in the transaction. Although there is no central, public marketplace for borrowing shares for short selling, short selling of stocks with a large market capitalization is not difficult, since these stocks are widely held and have high levels of institutional ownership.\(^{30}\) It is not as simple or inexpensive to borrow stocks that have a smaller market capitalization, are closely held, or are believed to be overvalued, since these stocks are highly sought for borrowing.\(^{31}\)

One important distinction between a short sale and a regular purchase of stock—when an investor purchases a stock in anticipation that its value will rise—is the “theoretically unlimited” risk of loss when an investor sells short.\(^{32}\) If an investor made a run-of-the-mill stock purchase to hold it for the long term (a “long” position), and the stock price later fell to zero, he would not lose any more than the amount he paid for that stock. But if the investor short sells a security (a “short” position) and the market price of that stock keeps rising, he continues to lose money because he is still required to replace the borrowed stock, and this purchase will now have to be made at a progressively higher price.\(^{33}\) When a stock with a limited supply is in high demand by the market, forcing prices

---


\(^{31}\) Id. at 8.


up, a “short squeeze” may occur.\(^{34}\) Thus, when short sellers rush to exit their positions, prices continue to increase, and losses to short sellers continue to mount.\(^{35}\) In addition to the risk of unlimited financial loss, if an investor is holding a short position and dividends are declared on his borrowed stock, he will be responsible for reimbursing the lender for the total amount of the dividends.\(^{36}\)

1. Reasons for the Negative Perception of Short Selling

Many observers see short selling as a “bet[] against the team,”\(^{37}\) anti-economic growth,\(^{38}\) or “un-American” since short sellers profit when they correctly bet that a stock’s price will fall.\(^{39}\) Consequently, in times of economic trouble and difficult world events, “shorts” are often looked upon as scapegoats,
particularly in the wake of large stock market declines.\footnote{Nasty, Brutish and Short; Short-Selling, ECONOMIST, June 21, 2008, at 46; see also Lamont, supra note 30, at 7.} During World War I, the New York Stock Exchange ("NYSE") imposed special restrictions on short sales because of fear that the Germans would use the practice to manipulate and drive down stock prices,\footnote{Id. at 8.} and because it would harm the morale of the market.\footnote{Id. at 8.} In the years following the stock market crash of 1929, when Congress began developing the extensive regulations of securities markets that are in place today, there was even a proposal to ban short selling altogether.\footnote{Macey et al., supra note 27, at 801.} Congress and others again called for increased scrutiny of short sales of equity stocks after the precipitous stock market decline in 1987.\footnote{Id. at 799.} Additionally, after the terrorist attacks in New York and Washington, D.C. on September 11, 2001, U.S. and European market regulators even investigated a spike in short sales of airline and insurance stocks in the days leading up to the attacks to determine if short sellers knew about the plan.\footnote{Cassell Bryan-Low, Initial Investigation Fails to Dig Up Evidence Linking Level of Short Selling to Terrorists, WALL ST. J., Sept. 24, 2001, at C2. The SEC investigated thirty-eight firms whose stock was identified as having unusual trading activity in the days before the terrorist attacks, but no connection was affirmed. Susanne Craig, SEC Examines Trading in Firms Before Sept. 11, WALL ST. J., Oct. 3, 2001, at C1.}

In addition to being criticized for profiting from falling stock prices, another reason for the vilification of short sellers is the fear that they aim to benefit from market manipulation.\footnote{Proposed Amendments to Regulation SHO and Rule 10a-1, Exchange Act Release 54,891, 71 Fed. Reg. 75,068, 75,070 (Dec. 13, 2006) [hereinafter 2006 Proposed Amendments to Regulation SHO and Rule 10a-1].} In a “bear raid,” investors continually short sell an equity stock in an attempt to influence “less informed” shareholders of a negative price outlook on the security, in the hopes they will sell off their shares.\footnote{Id. at 799.} This is problematic because the rapid decline of a stock price caused by market manipulation could prompt margin calls or liquidations.\footnote{Id.; see also Rob Curran & Geoffrey Rogow, Margin Calls, Redemptions Weigh on Market, WALL ST. J. ONLINE, Oct. 28, 2008, http://online.wsj.com/article/SB122513281830072753.html?mod=googlenews-wsj.} Many investors trade on the margin, meaning that they trade with borrowed money
while keeping a certain level of cash with brokers as collateral. Brokers and exchanges determine the level of cash necessary, based on risk. Generally, declining stock prices and increasing market volatility increase the risk of loss, prompting brokerages to ask for more collateral. This demand is known as a margin call. Often, to raise cash for margin calls, investors sell stock which can depress prices and create a cycle which results in another margin call. In addition, when investors continue to sell stocks to raise cash for margin calls, they are less likely to buy stocks, further depressing market prices.

The 2008 credit crisis illustrated this fear of short sellers manipulating the market when, once again, regulators singled out the shorts for contributing to financial chaos. After the failure of Bear Stearns in March 2008 there were accusations that short sellers spread rumors about companies to put downward pressure on stock prices, thus allowing investors with a short position to reap profits. These accusations eventually led the SEC to issue an emergency order in July 2008 announcing that it was going to begin investigating whether short sellers were colluding to manipulate the markets for their own gain.

The subsequent failure of Lehman Brothers and concerns about market volatility led the Commission to issue a total ban on the short sale of financial stocks a few months later. In September 2008, by installing a ban on the short sale

---

50 Curran & Rogow, supra note 48.
51 Id.
52 Id.
53 Id.
54 Jenny Anderson, A New Wave of Vilifying Short Sellers, N.Y. TIMES, Apr. 30, 2008, at C1. During a Senate hearing probing the failure of Bear Stearns, Senator Christopher Dodd suggested that there was collusion in the marketplace targeting the bank. Id.
55 Id.
57 Press Release 2008-211, supra note 8. The Chicago Board of Exchange Volatility Index (VIX) “is a key measure of market expectations of near-term volatility conveyed by S&P 500 stock index option prices.” Chicago Board of Exchange—Micro Site, Introduction to VIX Futures and Options, http://www.cboe.com/micro/vix/introduction.aspx (last visited Oct. 6, 2009). The VIX is considered a barometer of investor sentiment (it is often called the “fear index”) and uses options prices to estimate the range of movement of the S&P 500 for the following 30 days. Tom Lauricella & Aaron Lucchetti, Dow Slides Again, Down 514.45—S&P at a 5-Year Low; What’s Behind the Surge in the VIX ‘Fear’ Index?, WALL ST. J., Oct. 23, 2008. After Lehman declared bankruptcy on September 16, 2008, the VIX Index rose above 30 for
of financial stocks, the SEC sought to avert a “crisis of confidence” resulting from sharp declines in stock prices. The Commission posited that a drop in stock prices would potentially affect the “liquidity” and “ultimate viability” of financial institutions and damage the broader securities market. Though it soon became apparent that this ban was temporary, it provoked much negative reaction in the financial community. Regulators in other nations soon followed suit, hoping to prevent the volatile equities markets from spiraling further out of control.

Legislators are not the only ones who have criticized the practice of short selling. Companies have also taken actions against short sellers in a number of ways, including issuing stock with certain restrictions that make short selling impossible, taking legal action against short sellers, and reporting short sellers to regulatory agencies. At the same time since 2007, which was when the first news of a coming subprime crisis hit the markets. The VIX remained above 30, hitting a high of 81 in October 2008. For comparison, the VIX rose above 30 on only three trading days in 2007 (and went no higher than 31.09). VIX Daily Closing Prices 2004-Present, Chicago Board of Exchange, http://www.cboe.com/micro/vix/historical.aspx (click on link to VIX data for 2004 to present) (last visited Oct. 6, 2009). The previous high for the VIX was 52.05 on September 21, 2001, when fear was high following the September 11, 2001 terrorist attacks. Lauricella & Lucchetti, supra.

58 Sept. 2008 Emergency Order, supra note 8, at 55, 175.
59 Id.
60 See Scannell & Karmin, supra note 8.
62 See Kara Scannell, Short Sale Ban Spreads Around Globe, Sept. 22, 2008, at C3 (detailing implementation of short sale ban by Australia, the Netherlands and Taiwan.).
63 Lamont, supra note 30, at 3. In a 2003 study, Professor Lamont examined the long-term returns of 270 firms that attempted to actively discourage short selling through threats, legal action or accusations of improper activity. Id. at 4-5. In the year following these firms’ actions against short sellers, they had an average return compared to the overall stock market of -24%. Id. at 3. This indicates that the short sellers were correct and that the securities of the firms in question were overpriced. Id. An example of a company issuing securities creatively to prevent short selling occurred in 2006 when Pegasus Wireless granted a dividend to stockholders for every 10 shares held; the dividend was distributed as a stock warrant, but the company refused to issue the warrant unless investors held their shares in their own name. Jenny Anderson, A Bet Against Those Who Bet Against the Company, N.Y. TIMES, Sept. 1, 2006, at C7. Brokers often hold shares in their accounts (and in the broker’s name) for clients and then lend the shares held to short sellers. Id. This dividend structure was problematic because it forced brokers to recall the stock to prove shareholders’ identities, which put
time that the SEC announced the temporary ban on short selling of financial stocks in 2008, it also enacted a rule that required hedge funds and other institutional investors to disclose which companies’ securities they were holding short.\textsuperscript{64} This provoked an immediate backlash from the investment community over concerns that company executives would no longer provide information to investors who were known to be shorting their company’s stock.\textsuperscript{63} The SEC eventually backed down from this rule over the investors’ concerns.\textsuperscript{65}

2. The Positive Effect of Short Selling on the Markets

Despite the persistent animosity toward shorts during troubled economic times, short sales of securities do have several positive effects on stock markets. Two of the most important areas of the market that are positively affected by short selling are “pricing efficiency” and “liquidity.”\textsuperscript{67} If stock markets were perfectly efficient, short sales would not be necessary because all stocks would be correctly valued.\textsuperscript{68} However, since markets are in fact imperfect, short sales can help correct inefficiencies created by “asymmetric information, taxes, or other imperfections” by moving prices closer to equilibrium.\textsuperscript{69} This is because a short sale is an investor’s way of “inform[ing] the market of [his negative] evaluation of future a short squeeze on the stock. \textit{Id.} The CEO of Pegasus claimed that these actions were taken not to harm short sellers but because the company was concerned that phantom (fake) shares of its stock were being traded. \textit{Id.} Nonetheless, the company’s share price rose 30\% after the warrant plan was announced. \textit{Id.}


\textsuperscript{65} One prominent investor, Jim Chanos, called the SEC’s rule “akin to the government suddenly requiring Coca-Cola to disclose their secret formula for free to all their competitors.” Beth Healy & Ross Kerber, \textit{Short-sellers Cry Foul After Ban, BOSTON GLOBE}, Sept. 26, 2008, (internal quotation marks omitted).


\textsuperscript{67} 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, \textit{supra} note 46, at 75,069.

\textsuperscript{68} The efficient market theory posits that in an efficient market, all investors would have perfect information, thus investments would be properly valued and as a result, short selling would be unnecessary. \textit{See William W. Bratton, CORPORATE FINANCE} 15 (6th ed. 2008).

\textsuperscript{69} \textit{See Powers et al., supra} note 32, at 235-36.
Many feel that short sellers keep exuberance in check and help to quiet “noise traders.” Moreover, short sellers provide liquidity because not only must shorts cover their sales by buying stocks, but an investor may also be more willing to take a bigger risk on a long position if he can hedge himself with a short position. In 1931, Richard Whitney, then-President of the NYSE, testified before Congress that he believed the exchange would have been forced to close after the 1929 stock market crash if it were not for the willingness of short sellers, the only investors that made money when the NYSE crashed, to put their money back into the depressed market after stock prices dropped.

The perceived skepticism or pessimism of short sellers plays a very important role in the market. Short sellers are often among the first to detect corporate fraud and were among the first to issue warnings about ill-fated companies such as Enron and Tyco. Since many short sellers are among the most informed and sophisticated investors, they may perceive financial malfeasance before market regulators, despite the best attempts of the SEC and other regulatory agencies to keep tabs on companies. Individual investors that research the fundamentals of a company and its stock, operating only with an eye for profit, may simply have more resources and perhaps more motivation to delve into the details of a firm’s financial information when making an investment decision.

---

70 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,070.
71 Crisp, supra note 38, at 142. Noise traders act “for reasons generally unrelated to an accurate measure of an asset’s fundamental value. . . . and might act on market momentum, misinformation or poor strategy.” Id. at 141 (citation omitted). An overly positive view on a stock is less likely to be challenged than a view that is excessively negative. See Powers et al., supra note 32, at 241.
72 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,069.
73 See Macey et al., supra note 27, at 800.
74 Id. at 801-02.
75 See Anderson, supra note 54. Jim Chanos, who runs one of the largest short funds in the world for Kynikos Associates, calls short sellers “financial detectives.” Nasty, Brutish and Short, supra note 40, at 46 (internal quotation marks omitted). Coincidentally, Mr. Chanos is also credited for being one of the first skeptics of Enron’s financial statements. Id.
76 See Richard Sauer, Bring on the Bears, N.Y. TIMES, Oct. 6, 2006, at A25 (“By putting their money where their mouths are, short sellers are the only market participants with an incentive to deflate bubbles and inject pessimistic information into the market.”).
B. The History of the Uptick Rule—1938-2007

Considering the historic sentiment against short selling, it is not surprising that following the 1929 market crash, short selling was a concern to legislators. However, the Securities Exchange Act of 1934 (“Exchange Act”) did not specifically govern short sales; it instead delegated regulation of short selling to the SEC. In 1934, the newly-created SEC released a list of rule recommendations for adoption by national exchanges. One of the sixteen rules recommended was to prohibit the short sale of a security at a price below the last previous sale price of that security. The SEC believed this formulation of the short selling rule would prevent abusive short selling on exchanges while also “preserv[ing] those features of short selling which are in the public interest.” Thus, the rule would protect the investing public while enabling markets to operate unhindered, allowing them to reap the inherent benefits of short selling.

After a decline in the market in the fall of 1937, the SEC undertook a study to examine whether short selling exacerbated the drop in stock prices. Though several studies of short selling were still under way at the time, the SEC released some data publicly in January 1938 that suggested that short selling increased in a declining market and that in such a

77 Macey et al., supra note 27, at 801-03.
78 Id. at 802-03. Regulatory power over short selling is codified at 15 U.S.C. § 78j (a)(1) (2006). This provision states:

It shall be unlawful for any person, directly or indirectly, by the use of any means or instrumentality of interstate commerce or of the mails, or of any facility of any national securities exchange – (a) (1) To effect a short sale, or to use or employ any stop-loss order in connection with the purchase or sale, of any security registered on a national securities exchange, in contravention of such rules and regulations as the Commission may prescribe as necessary or appropriate in the public interest or for the protection of investors.

Id.
79 See FIRST ANNUAL REPORT OF THE SEC, supra note 16, at 40-44.
80 OFF. OF ECON. ANALYSIS, SEC, ECONOMIC ANALYSIS OF THE SHORT SALE PRICE RESTRICTIONS UNDER REGULATION SHO PILOT 12 (2007), http://www.sec.gov/news/studies/2007/regshopilot020607.pdf [hereinafter OEA ECONOMIC ANALYSIS]. The SEC recommended as its Sixteenth Rule governing exchanges that “[n]o member shall use any facility of the exchange to effect on the exchange a short sale of any security in the unit of trading at a price below the last sale price of such security on the exchange,” but the rule was not implemented at that time. FIRST ANNUAL REPORT OF THE SEC, supra note 13, at 44.
market, “short sales are seriously destructive of stability.” It was based on this limited information that the SEC concluded that short selling needed to be further regulated, and it adopted a set of rules, effective February 8, 1938, which attempted to “prohibit short selling in a declining market.” These rules included the uptick rule, previously recommended for adoption by the exchanges, as well as a rule that all sell orders be marked “short” or “long.” It is notable that when the SEC implemented this rule change, it made clear that it wished to formulate the short selling rules in such a way as to “avoid placing undue burden or inconvenience on transactions,” and that it would revisit the necessity of these rules if they were deemed unnecessary by the Commission’s ongoing studies or if they had a negative impact on the market.

The 1938 version of the uptick rule went virtually unchanged until the SEC removed all price test rules for short sales in July 2007. In 1939, the SEC modified the rule slightly to allow short selling on a zero uptick, making a short sale permissible if the sale occurred at the same price as the last trade, if the price of the next-to-last trade was an uptick.

---

83 Id. at *5.
84 Id. at *1.
85 Id. at *2 (Rule X-10A-1(a)). A tick refers to a move upward or downward in the price of a security.
86 Id. (Rule X-10A-1(b)).
87 Id. at *1. The 1937 study has been criticized as inadequate due to its limited scope and the short time frame for which it was performed. Short Sales of Securities, Exchange Act Release No. 13,091, 41 Fed. Reg. 56,530, 56,533 (Dec. 28, 1976). In addition, after the tick rule was implemented, no further data compiled from the 1937 study was publicly released despite being promised in the initial report. Id.
88 The “downtick” and “uptick” distinction is a matter of semantics. The early (downtick) version of the rule prevented a short sale when the last price movement was downward, while the later (uptick) version of the rule prohibits a short sale at a price that is not at a tick price higher than the last previous trade. See Exchange Act Release No. 1548, 1938 WL 32911, supra note 82, at *1 (“No person shall . . . effect a short sale of any security at or below the price at which the last sale thereof, regular way, was effected on such exchange.”).
89 An illustration of the zero plus tick test and compliance with Rule 10a-1 in a sale sequence follows:

Last sale: 47
Next sale: 47.04 — Plus tick compared to last trade; short sale permitted
Next sale: 47.04 — Zero-plus tick compared to next-to-last trade; short sale permitted
Next Sale: 47.00 — Minus tick compared to next-to-last trade; no short sale permitted
SEC felt that this rule met three important objectives: (1) it did not unfairly restrict short sales when the market was increasing; (2) it prevented bear raiders from driving down prices in a declining market; and (3) it prevented short sellers from using all bids at one price level, thereby causing long sellers to set progressively lower prices. The uptick rule applied to all securities registered on national securities exchanges and also regulated trades of securities “admitted to unlisted trading privileges” on national securities exchanges, if the trades were reported in accordance with an “effective transaction reporting plan.” However, the uptick rule did not apply to over-the-counter bulletin board securities or pink sheets, as neither of these types of securities are traded on a national exchange.

The short sale price test rule operated slightly differently as applied to securities trading on the National Association of Securities Dealers Global Market (NASDAQ). Before becoming a national securities exchange, the NASDAQ operated a price test for short selling called a bid test. The bid test did not allow short sales at or below the current highest bid when that bid was less than the previous highest bid. When the NASDAQ applied to the SEC to become a national exchange in January 2006, it requested, and received, an

Next Sale: 47.00 — Zero-minus tick compared to next to last trade; no short sale permitted

2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,070.

See id. (citation omitted); Macey et al., supra note 27, at 803-04 (citation omitted).


2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,070. Over-the-counter bulletin board securities (OTCBB) are generally small, risky, and traded infrequently. Investopedia, Over-The-Counter Bulletin Board (OTCBB), http://www.investopedia.com/terms/o/otcbb.asp (last visited November 5, 2008). There are no listing requirements to trade on the OTCBB, though companies must file financial statements with the SEC. Id. Pink sheets are securities that are not traded on an exchange, do not have listing requirements and are not required to file with the SEC. Investopedia, Pink Sheets, http://www.investopedia.com/terms/p/pinksheets.asp (last visited Oct. 6, 2009).

2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,070-71.

Id. The SEC granted temporary approval in 1994 for the NASD to use this bid test (former NASD Rule 3350). Id.

Id. at 75,071 & n.29.
exemption from the uptick rule.\footnote{Id. at 75,071.} The bid test remained in place until the repeal of all short sale price tests in 2007.\footnote{Id. The SEC granted this exemption in large part due to the fact it was in the process of conducting a Pilot to study the effect of removing a short sale price test rule. Id.; see also infra Part III. The Commission did not want to jeopardize the quality of the pilot data or impose costs on the NASDAQ to implement a rule when it was possible the rule would be temporary. 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,071. After the NASDAQ was accepted as a national exchange, the bid test was codified as NASD Rule 5100. At this time, the SEC also exempted NASDAQ securities, newly listed on a national exchange, from the bid test when traded on non-national exchanges. Id.} Though the NASDAQ-listed securities were technically exempt from the uptick rule, these securities were still regulated by a price test.

To enforce the uptick rule the SEC used a marking requirement. When stock trades were executed, each order placed with a broker-dealer had to be marked as “short” or “long.”\footnote{See Regulation of Short Sales, 17 C.F.R. § 242.200(g) (2007). If the sale of a stock was subject to a particular exemption from the uptick rule, the order was required to be marked “short exempt.” Id. § 242.200(g)(2).} As such, if an investor mismarked a trade as a purchase of shares for a long position, but actually purchased a short position, and did not observe the price test requirement of the uptick rule, the investor would be found in violation of Rule 10a-1. For example, the SEC brought an enforcement action against Sandell Asset Management after the firm began short selling shares of Hibernia Corporation, a New Orleans bank holding company, immediately after Hurricane Katrina.\footnote{Press Release 2007-216, SEC, SEC Charges New York Hedge Fund Adviser With Short Sale Violations in Connection With Hibernia-Capital One Merger (Oct. 10, 2007), http://sec.gov/news/press/2007/2007-216.htm. Sandell Asset Management settled with the SEC for approximately $8 million without admitting or denying the charges. Id.} Sandell “held a large long position in Hibernia.”\footnote{Id.} The firm was apparently concerned that the natural disaster would decrease the offer price for Hibernia in a pending acquisition of the bank by another company.\footnote{Id.} To protect Sandell against potential losses if the acquisition deal fell through, Sandell employees allegedly tried to short sell as many shares as possible to hedge its Hibernia investment and in the process, “short” sale orders were falsely marked as “long” sale orders.\footnote{Id.}

Until its repeal in 2007, the uptick rule was in place as a backstop to regulate harmful or manipulative short selling in a declining market for nearly as long as the existence of the
SEC itself. However, as a result of market modernization and a growing list of exceptions to the rule, the SEC sought to revisit whether the rule needed to be modified to fit the trading practices of the twenty-first century.

III. THE REPEAL OF THE UPTICK RULE AND RENEWED CONSIDERATION OF A PRICE TEST RULE

The uptick rule has been considered controversial for decades. During the seventy years the uptick rule was in effect, the SEC studied its efficacy and necessity on several occasions, but it made no significant changes to the rule until its repeal in 2007. In addition to studying the effects of short selling on the market crash in fall 1938, another extensive study was performed in 1962. Between 1939 and 1963, the NYSE lobbied unsuccessfully for the SEC to change the price test rule and allow short sales of a security at any price, so long as that price was higher than the closing price of the security on the previous trading day. The necessity of this rule was examined again in 1976. Given the history of the uptick rule, it was no surprise that when the SEC issued a Concept Release in 1999 in order to garner public comments on the short selling rules in an effort to “modernize” regulation of short selling, it received more than 2700 comment letters.

A. The 2007 Repeal of the Uptick Rule

The SEC felt it was necessary to re-examine the short selling rules due to several market developments, including the

---

105 See id. at 62,973.
106 Short Sales of Securities, Exchange Act Release No. 13,091, supra note 87, at 56,534. The Special Study, published in 1963, was more extensive than the 1937 study and concluded that while short sales as a percentage of total market volume increased in a declining market, the tick rule should be accompanied by a rule to “cope” with short selling during market declines because plus and zero plus ticks could occur in “sharply declining markets.” Id.
107 The uptick rule only applied to transactions on national exchanges regulated by the SEC (such as the New York Stock Exchange) and to securities traded on national exchanges. Thus the rule did not apply to over the counter (OTC) sales of securities that are not traded on an exchange. See David C. Worley, The Regulation of Short Sales: The Long and Short of It, 55 BROOK. L. REV. 1255, 1261 (1990).
increase in NASDAQ-listed securities trading off that market, the impact of electronic trading and decimalization, and the effect of the now-commonplace practice of trading of options.\footnote{110} There was also speculation that hedge funds heavily pressured the SEC to re-evaluate the short selling regulations.\footnote{111}

In 2003, after receiving and examining comments on its 1999 Concept Release, the SEC submitted Regulation SHO for public consideration.\footnote{112} Rather than rescind the uptick rule immediately, Regulation SHO proposed a pilot test period (the “Pilot”) during which time the uptick rule would be inapplicable to certain stocks.\footnote{113} This would allow the SEC to obtain and study information about the trading activity for stocks not subject to a short sale price test.\footnote{114} After an initial delay, the Pilot began in May 2005.\footnote{115} Although the test was supposed to end in April 2006, the SEC extended the Pilot until August 2007 to give the Commission enough time to evaluate the data and determine whether to modify or repeal the price test rules in place.\footnote{116} The Pilot exempted approximately 1000 stocks, chosen from the Russell 3000 Index, from the uptick rule.\footnote{117} The SEC selected the Pilot stocks using a methodology that it felt would “give[d] due consideration to the liquidity, volatility, market depth and trading market of these securities.”\footnote{118}

\footnote{110} See id.
\footnote{111} See Jeff Benjamin, Did Repeal of the Uptick Rule Unleash Market Havoc? Surge of Volatility, Rising Number of Short Sales Cited as Evidence, 11 INV. NEWS 3, 3 (2007).
\footnote{113} Id. at 62,983.
\footnote{114} Id.
\footnote{116} Id. The primary reason for the extension was to prevent securities markets from having to make costly modifications to their systems and procedures more than once in the event the Commission decided to repeal a short sale price test rule. Id. The end of the Pilot coincided with the expiration of the temporary order suspending price tests. Id.
\footnote{118} Order Suspending the Operation of Short Sale Price Provisions, 69 Fed. Reg. at 48,032. The securities included in the Pilot were chosen from the Russell 3000 as of June 25, 2004 and the test group included only those stocks subject to Rule 10-(a)(1), which were all those traded on the NASDAQ and those listed on the NYSE or American Stock Exchange (Amex). Id. The stocks were grouped by the three exchanges
Following the conclusion of the first year of the Pilot period, the SEC Office of Economic Analysis ("OEA") analyzed the Pilot data and issued a report in early 2007 to assist the Commission in deciding whether to repeal the price test rule, to install an alternative price test, or to retain the price test already in place. Based on the results of the OEA's analysis, a public roundtable discussing the OEA report, and four additional studies performed by independent parties, the SEC concluded that "[g]enerally, the Pilot Results supported removal of current price test restrictions." The Pilot data revealed little relationship between "manipulative short selling" and the restrictions imposed by the uptick rule.

Another factor that contributed to the SEC's decision to reconsider and eventually repeal the uptick rule was the Commission's goal of creating a "more consistent regulatory environment for short selling." During the nearly seventy years the uptick rule was in effect, the SEC granted numerous exceptions to the rule, primarily as a result of the modernization of the markets and the evolution of trading practices. Generally, the SEC allowed statutory or written exemptions for types of transactions that the uptick rule was not designed to prevent or for activities that were not deemed abusive.

and then ranked by "average daily dollar volume" during the preceding year. Id. Every third stock (beginning with the second stock on the list), was then chosen in order to have "a more representative daily dollar volume sample." Id at 48,032, n.7. 50% of the test stocks were listed on NYSE, 2.2% on the Amex, and 47.8% on the NNM. Id. at 48,032. The approximately 2000 stocks in the index not chosen for the Pilot constituted the control group and the percentage distribution across the three exchanges were nearly identical. Id.

See Regulation SHO and Rule 10a-1 Final Rule Release, supra note 9, at 36,349, n.17.

Id. at 36,349. See generally OEA ECONOMIC ANALYSIS, supra note 80. Though much of the economic analysis in this report is beyond the legal scope of this Note, the conclusions of the report that the SEC utilized in determining to rescind the uptick rule will be discussed in this Part and infra Part V.

2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,073.

Id. at 75,068.

Id. at 75,071. The statutory exceptions to the uptick rule can be found in section e of Rule 10a-1. 17 C.F.R. § 240.10a-1(e)(1)-(12) (2007). The SEC may also exempt transactions from the uptick rule upon "written request." See 17. C.F.R. § 240.10a-1(f) ("This rule shall not prohibit any transaction or transaction which the Commission, upon written request or upon its own motion, exempts, either unconditionally or on specified terms and conditions.").

granted to allow market makers and specialists to sell short on a zero minus tick, to ensure maintenance of prices and market liquidity. In addition to these exceptions, in 2000-2001, U.S. securities markets implemented decimalization, which changed the pricing of shares from 1/16th minimum increments to $.01 increments. In a decimalized trading environment, one tick of a stock price now had a value of $.01 as compared to the old minimum tick of 1/16th of one dollar, or $.0625. In repealing the rule, the SEC and other critics of the uptick rule argued that decimalization made the rule less effective since a penny tick test would be less effective at slowing down short sellers.

The SEC also felt repealing the uptick rule was necessary to prevent exchange arbitrage. This is because regulatory differences between exchanges could put exchanges with a price test rule at a competitive disadvantage to exchanges without such a rule. For example, in 1985, the NYSE publicized that it wanted the price test rule relaxed in order to combat competition from the London Stock Exchange, an exchange without a short sale price test. As financial markets have continued to globalize over the past twenty years, this competition still exists. The results of the Pilot lent support to these concerns, as it indicated that short selling as a portion of total volume increased by 2% in the absence of price test restrictions. The Commission construed this data to mean that investors could be more inclined to conduct their short sale transactions at market centers without short sale price tests.

125 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,072.
127 See April 2009 Amendments to Regulation SHO, supra note 4, at 18,061 (“[T]he Commission noted [at the time of repeal of the uptick rule] that decimal increments had resulted in a rule that was no longer suited to the wide variety of trading strategies and systems used in the marketplace.”).
128 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,075.
129 Macey et al., supra note 27, at 804.
130 OEA PILOT STUDY, supra note 80, at 35.
131 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,075; see also OEA PILOT STUDY, supra note 80, at 35-36.
Whether hedge funds influenced the SEC’s decision to remove price test restrictions on short selling is debatable.¹³² But the timing of the Pilot test and the repeal of the uptick rule certainly coincided with the ascension of hedge funds.¹³³ While there is no agreed-upon, clear-cut definition of a hedge fund,¹³⁴ managers of these private investment funds often rely heavily on short selling as an investment strategy.¹³⁵ Traditional investment companies, which are regulated by the Investment Company Act,¹³⁶ are subject to stringent rules requiring extensive disclosure of short sales in the companies’ investor prospectuses and annual reports.¹³⁷ Because hedge funds are not subject to the same regulation, the growth of hedge funds likely increased the visibility and incidence of short selling since they may freely use short selling strategies to maximize returns.¹³⁸ Predictably, hedge funds were publicly supportive of both the SEC’s initial 2003 proposal suggesting the Pilot study,¹³⁹ and the 2006 proposal to repeal the uptick rule.¹⁴⁰

¹³² See Benjamin, supra note 111, at 3; Editorial, Opposing Uptick Rule is Truly Short-Sighted, INVESTOR’S BUS. DAILY, Oct. 16, 2008, at A10.
¹³³ Hedge fund employees are famously secretive about their investments and strategies, partly due to the competitive nature of the industry and partly because they raise funds privately, without advertising or public solicitation. Jenny Anderson, Hedge Funds Walk a Hard Line Between Silence and Sharing, N.Y. TIMES, Feb. 9, 2007, at C7. Private fundraising means that hedge funds fall outside the scope of the Securities Act of 1933. Id. This secrecy and the fact that there are no public reporting requirements make it difficult to estimate the precise increase in the amount of assets under management by hedge funds. However, for the sake of perspective, one estimate approximates that the number of global assets under management in hedge funds increased from “$50 billion in 1990 to approximately $1 trillion at the end of 2004.” Burton G. Malkiel & Atanu Saha, Hedge Funds: Risk and Return, 61 FIN. ANALYSTS J. 80, 80 (2005). In 2003, the “significant growth” of hedge funds prompted the SEC to compile a report to study implications of this growth given the “lack of information” available about hedge funds. SEC, STAFF REPORT TO THE SEC. EXCH. COMM’N, IMPLICATION OF THE GROWTH OF HEDGE FUNDS vii (2003), available at http://www.sec.gov/news/studies/hedgefunds0903.pdf.
¹³⁴ In a 2003 Staff Report, the SEC provided its general definition, calling a hedge fund “an entity that holds a pool of securities and perhaps other assets, whose interests are not sold in a registered public offering and which is not registered as an investment company under the Investment Company Act.” SEC, STAFF REPORT TO THE SEC. EXCH. COMM’N, supra note 133, at 3.
¹³⁵ See id. at 5.
¹³⁷ SEC, STAFF REPORT TO THE SEC. EXCH. COMM’N, supra note 133, at 43, n.147.
¹³⁸ Id. at 42-43.
¹³⁹ See Letter from John G. Gaine, President, Managed Funds Assoc., to Jonathan G. Katz, Sec’y, SEC (Jan. 26, 2004) (http://www.sec.gov/rules/proposed/ s72303/managedfunds012604.htm). On behalf of the Managed Funds Association, an organization which represents an industry group of alternative investment fund professionals (including those employed by some of the largest hedge funds), Mr. Gaine
However, despite initially strongly supporting the SEC’s decision to repeal the uptick rule, some hedge funds lobbied for the reinstatement of the rule once the SEC imposed the emergency ban on short sales in September 2008. These calls by hedge funds to reinstate the uptick rule came fast and furious after the SEC banned short sales and began requiring funds to disclose short positions. This outcry over the disclosure requirement was part of the reason the SEC softened its stance and modified the regulations slightly so that only short positions in excess of a fair market value of $10 million had to be reported (as opposed to the initial requirement of $1 million).

In 2006, the SEC announced a proposal to repeal the uptick rule and solicited another round of public comments, most of which were in favor of the rule change. The comments submitted were mostly in line with the SEC’s view expressed in the proposal, pointing out that improvement in market surveillance and transparency rendered the backstop of a price test rule unnecessary. The comments asserted that the elimination of price test restrictions would allow the market to benefit from the merits of short selling, such as pricing efficiency and liquidity, while eliminating investors’ operational costs directly associated with compliance with the rule. Interestingly, two individual investors urged the SEC to keep the uptick rule in place to prevent bear raids. Another comment letter from a finance professor agreed vigorously with the SEC’s decision, but also made an interesting point that perhaps if more short sale data were publicly available, shorts

“encourage[d] the Commission to . . . move expeditiously toward the complete removal of short sale price regulation.” Id.

140 See Letter from John G. Gaine, President, Managed Funds Assoc., to Nancy M. Morris, Sec’y, SEC (Feb. 12, 2007) (http://www.sec.gov/comments/s7-21-06/s72106-31.pdf) (urging the SEC to repeal all price test restrictions and applauding its efforts “towards removing obsolete and unnecessary regulations”).


142 See Healy & Kerber, supra note 65.


144 Regulation SHO and Rule 10a-1 Final Rule Release, supra note 9, at 36,350.

145 Id.

146 Id.

147 Id.
would not always be the first parties blamed during market declines.\footnote{Letter from James J. Angel, Assoc. Prof. of Fin., Georgetown Univ., to Nancy M. Morris, Sec’y, S.E.C (Feb. 14, 2007), http://www.sec.gov/comments/s7-21-06/s72106-35.pdf.}

One comment letter the NYSE submitted following the announcement of the proposal to repeal the uptick rule was particularly notable. The NYSE expressed concern that the entire Pilot test took place during a period in which the market was relatively stable.\footnote{See Letter from Mary Yeager, Assistant Sec’y, N.Y. Stock Exch., to Nancy M. Morris, Sec’y, SEC (Feb. 14, 2007), http://www.sec.gov/comments/s7-21-06/s72106-34.pdf.} As such, the NYSE noted that removal of the rule during a period of “unusually rapid and large market decline” could not be measured.\footnote{Regulation SHO and Rule 10a-1 Final Rule Release, supra note 9, at 36,350; see also Letter from Mary Yeager, Assistant Sec’y, N.Y. Stock Exch., to Nancy M. Morris, supra note 149.} Conversely, when the SEC established the rule in 1938, it did so in part based on a study of two one week periods in September and October of 1938 that were “characterized by a large volume of trading, erratic intermediate price movements and intensive liquidation.”\footnote{Exchange Act Release No. 1548, 1938 WL 32911, supra note 82, at *1.} The NYSE also expressed its belief that national exchanges should have the option to suggest price-testing rules in unstable markets.\footnote{Letter from Mary Yeager, Assistant Sec’y, N.Y. Stock Exch., to Nancy M. Morris, supra note 149.} Not surprisingly, immediately following the SEC’s ban on short selling of financial stocks in 2008, the Chief Executive of the NYSE, Duncan Niederauer, publicly announced that he favored the return of the uptick rule, especially in volatile market conditions.\footnote{Geoffrey Rogow, NYSE Chief Leans Toward Uptick Rule, WALL ST. J., Oct. 2, 2008, at C5.}

\section*{B. The SEC’s Reconsideration of the Uptick Rule}

Following the financial upheaval in 2008 and the appointment and confirmation of a new SEC Chairman in 2009, revisiting the regulation of short selling was an immediate priority for the SEC, due to the “extreme market conditions” and “deterioration in investor confidence.”\footnote{April 2009 Amendments to Regulation SHO, supra note 4, at 18,043.} In April 2009, the SEC sought comment on its proposal of two different approaches to regulate short selling.\footnote{See id. at 18,042.} The extensive
proposal detailed the approaches and included nearly two hundred questions for commenters to consider related to the proposed rules when providing feedback.\textsuperscript{156} The SEC received approximately four thousand comment letters regarding these proposals.\textsuperscript{157} Predictably, there was a wide range of responses, with many institutional commenters expressing concern that a short sale price test would have a deleterious effect on the efficiency and liquidity of the market,\textsuperscript{158} and many others urging the reinstatement of a price test regulation.\textsuperscript{159} This volume of responses helped prompt the announcement of a second public comment period in August 2009, proposing an additional version of a price test rule.\textsuperscript{160} In connection with the proposed rule releases, the SEC held a roundtable to discuss short sale price test regulation in May 2009 with various industry professionals.\textsuperscript{161}

The SEC suggested two different regulatory schemes in April 2009, the market-wide approach, with a permanent rule regulating short selling, and the circuit breaker approach that would implement a short selling regulation once the price of a security dropped precipitously, and could operate in conjunction with a market-wide or price test rule, or stand alone.\textsuperscript{162} As to the market-wide approach, the SEC solicited comments on two different price test rules, the “proposed uptick rule,” similar to the repealed Rule 10a-1 and the “proposed modified uptick rule,” similar to the bid test used by

\textsuperscript{156} See generally id.


\textsuperscript{159} See, e.g., Letter from Edward D. Herlihy and Theodore A. Levine, Wachtell, Lipton, Rosen & Katz, to Elizabeth Murphy, Sec’y, SEC (June 17, 2009), available at http://www.sec.gov/comments/s7-08-09/s70809-3690.pdf; E-mail from Glen Shipway to SEC, (June, 19, 2009) (http://www.sec.gov/comments/s7-08-09/s70809-3795.pdf).

\textsuperscript{160} August 2009 Amendments to Regulation SHO, supra note 157, at 42,033.


\textsuperscript{162} April 2009 Amendments to Regulation SHO, supra note 4, at 18,043.
NASDAQ before it became a national exchange. The Commission also suggested two alternative rules for the circuit breaker approach. The first is a “proposed circuit breaker halt rule” that would be “triggered by a severe price decline” in a stock and prohibit any short selling of that security. The second is a “circuit breaker price test rule[]” that would take effect when the price of a stock declined, while still allowing investors to short sell the security. Additionally, two alternative circuit breaker price test rules, the “proposed circuit breaker uptick rule” and the “proposed circuit breaker modified uptick rule,” were also suggested, paralleling the price test rules discussed under the market-wide approach. The SEC also discussed an “alternative uptick rule” in April 2009, but did not seek formal comment on this proposed regulation until August 2009. The alternative uptick rule is a price test rule that could be utilized in either a market-wide approach or a circuit breaker approach and would allow short selling of a stock only “at a price above the current national best bid.”

Before determining how to act, the SEC stated that it would also evaluate empirical data as it became available. The OEA had provided the Commission some preliminary data analyzing how a short sale price test would have affected the markets and whether short selling created downward pressure on stock prices during September 2008, at the height of the credit crisis. The SEC’s April 2009 proposal noted that the requests it received urging reinstatement of the uptick rule had not included any empirical data in support of these requests, but that it was “looking forward to receiving analysis of relevant data” related to the market effects of a price test rule, and the “costs and benefits of reinstating” some type of price test or circuit breaker rule.

In late February 2010, the SEC announced that it voted to adopt an alternative uptick rule that would take effect if a

---

163 Id. (internal quotation marks omitted); see also notes 88-89, 94-96 and accompanying text.
164 April 2009 Amendments to Regulation SHO, supra note 4, at 18,043.
165 Id.
166 Id.
167 Id.
168 August 2009 Amendments to Regulation SHO, supra note 157, at 42,033.
169 Id.
170 See April 2009 Amendments to Regulation SHO, supra note 4, at 18,049.
171 See id.
172 Id.
stock traded down more than ten percent in a day.\textsuperscript{173} Once the ten percent decline threshold is reached, short selling may occur only at a price above the best bid.\textsuperscript{174} During both of the 2009 comment periods, the SEC received a plethora of empirical data, but it did not point to any particular study that it found to be especially persuasive and conclusive when it implemented the new rule.\textsuperscript{175} This new regulation, Rule 201, will have broad coverage and “generally cover all securities . . . listed on a national exchange.”\textsuperscript{176} Once the price of a security declines 10%, the trading limits will continue in effect for the rest of the trading day as well as the following day.\textsuperscript{177} In enacting Rule 201, the SEC aims to regulate manipulative short selling and maintain investor confidence without “hav[ing] any negative effect on market liquidity and price efficiency.”\textsuperscript{178}

IV. REGULATION OF SHORT SELLING IN THE ABSENCE OF A PRICE TEST RULE

Upon announcing its decision to repeal the uptick rule in 2007, the SEC made sure to point to other statutes and regulations that enable the agency to police abusive short selling practices in the absence of a price test rule.\textsuperscript{179} Even without the uptick rule, it is still illegal to short sell stocks in contravention of the other securities rules and regulations.\textsuperscript{180} These statutes and rules were likewise enacted to protect investors and to maintain stable markets, and are the same regulations the agency would have used to enforce price manipulation and fraud through short selling even if the uptick rule were in place.

\footnotesize
\begin{itemize}
\item[173] See Floyd Norris, S.E.C. Restricts Short-Selling and Addresses a Global Accounting Shift, N.Y. TIMES, Feb. 25, 2010, at B3.\hfill
\item[174] See \textit{id.; see also supra note 23 and accompanying text.}\hfill
\item[175] 2010 Amendments to Regulation SHO, supra note 21, at 11,241-44.\hfill
\item[176] \textit{Id.} at 11,245. Options are not covered by Rule 201. \textit{Id.}\hfill
\item[177] \textit{Id.} at 11,244.\hfill
\item[178] \textit{Id.} at 11,248.\hfill
\item[179] 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,069 n.16. In addition to the laws discussed in this Part, the SEC also regulates short sales in connection with public securities offerings in Regulation M, Rule 105. Regulation M, 17 C.F.R. § 242.105 (2009). Public offerings have specific regulations that differ from rules that govern day-to-day trading activities. See Regulation M, 17 C.F.R. § 242.100-105 (2009).\hfill
\item[180] See supra note 78.\hfill
\end{itemize}
The Securities Act of 1933 (the “Securities Act”) section 17(a) is an anti-fraud provision that prohibits the use of interstate commerce to effect “fraud or deceit” through the sale of securities. A violation of subsection (a)(1) of this provision requires scienter. The Exchange Act also has two provisions that can be utilized to regulate abusive short selling. The first, section 9(a), prohibits the manipulation of securities prices, although this section applies only to securities listed on an exchange. The second provision, section 10(b), prohibits the use of any “manipulative or deceptive device[s]” in connection with the purchase or sale of securities and applies to any security, whether exchange-listed or not. Section 10(b) of the Exchange Act is extremely important in SEC enforcement because it is the general “catch-all” regulation that the SEC uses to implement needed rules to protect the investing public. The SEC enforces section 10(b) through Rule 10b-5. The activities proscribed under Rule 10b-5 are similar to those proscribed under section 17(a) of the Securities Act, but Rule 10b-5 is the farthest reaching anti-fraud rule promulgated by the Exchange Act, since it applies to any security.

Section 9(a) of the Exchange Act is the most important provision prohibiting price manipulation of securities listed on a national exchange. Since the securities of the largest and most frequently-traded companies are listed on national exchanges like the NYSE and the NASDAQ, this provision has bite. Several subsections of section 9(a) can be applied to manipulative short selling. The primary anti-manipulation

---

185 Joseph I. Goldstein et al., An Overview of Market Manipulation: Legal and Practical Aspects, in SECOND ANNUAL MARKET MANIPULATION 99, 105 (Joseph I. Goldstein et al. eds., 1990) [hereinafter An Overview of Market Manipulation] (“§ 10(b) is a broad ‘catch-all’ provision that empowers the Commission to prescribe rules that it deems necessary and appropriate to protect investors and the public interest.” (citing Ernst & Ernst v. Hochfelder, 425 U.S. 185, 203 (1976))).
186 Id. at 105.
187 Id. at 103.
188 Section 9(a)(1) primarily applies to “wash sales” and “matched orders.” Id. at 104. A wash sale is a transaction “which involves no change in the beneficial ownership” of a security. Securities Exchange Act of 1934 § 9(a)(1)(A), 15 U.S.C. § 78j(a)(1)(A) (2006). A matched order occurs when a purchaser (or seller) of a security enters an order for the purchase (or sale) of that security, when he knows that another party will be ordering a sale (or purchase) in the same security “at substantially the
provision is subsection 9(a)(2). To establish a violation of this provision, the SEC or the plaintiff must establish that: 1) a person made “a series of transactions in any security”; 2) that those transactions resulted either “in actual or apparent active trading in such security” or in a rise or decline in the price of such security; and 3) that the transactions were made “for the purpose of inducing the purchase or sale of such security by others.” Section 9(a)(2) also requires the plaintiff to establish manipulative intent. A would-be violation of section 9(a)(2) that is effected with a security not listed on an exchange is a violation of section 17(a) of the Securities Act.

Other provisions of section 9(a) that are relevant to the regulation of short selling include subsections 9(a)(3) and 9(a)(4), both of which apply to broker-dealers, as well as anyone else trading (purchasing or selling) stocks. Subsection 9(a)(3) prohibits any trader of a listed security from “inducing the purchase or sale” of a security by “circulating or disseminating information that market activity may occur that will cause the security’s price to rise or fall.”

Manipulation of securities prices is a “term of art” that is defined in different ways. Ernst & Ernst v. Hochfelder, 425 U.S. 185, 199 (1976) (“[Manipulation] connotes intentional or willful conduct designed to deceive or defraud investors by controlling or artificially affecting the price of securities.” (citation omitted)); see also In re Pagel, Inc., Exchange Act Release No. 22,280, 33 SEC Docket 1003, 1985 WL 548387, at *3 (1985) (“In essence, a manipulation is intentional interference with the free forces of supply and demand.” (citation omitted)).


A plaintiff in a private cause of action would also be required to prove that he relied on the transactions in question and that the transactions affected the plaintiff’s purchase or selling price. See Chemetron Corp. v. Business Funds, 682 F.2d 1149, 1164 (5th Cir. 1982), vacated on other grounds, 460 U.S. 1007 (1983).

In re Sharon M. Graham, Initial Decision Release No. 82, 1995 SEC Lexis 3457, at *26 (Dec. 28, 1995) (“Section[] . . . 9(a)(2) require[s] that the proscribed activities be engaged in with the requisite manipulative intent.”).

SEC v. Resch-Cassin & Co., 362 F. Supp. 964, 975 (S.D.N.Y. 1973) (“It is well settled that the manipulative activities expressly prohibited by § 9(a)(2) of the Exchange Act with respect to a listed security are also violations of § 17(a) of the Securities Act and § 10(b) of the Exchange Act when the same activities are conducted with respect to an over-the-counter security.”). For background on over-the-counter securities, see supra note 93.


buyers and sellers of listed securities from making “false or misleading” statements regarding any material fact related to a security to encourage a purchase or sale of that security, when the buyer or seller “had reasonable ground to believe” such statement to be “false or misleading.”

Section 10(b), the main anti-fraud provision of the Exchange Act, and Rule 10b-5, promulgated thereunder, are commonly used enforcement tools. Section 10(b) prohibits any person, by any means, from the “use . . . in connection with the purchase or sale of any security . . . [of] any manipulative or deceptive device.” Further, Rule 10b-5 makes it unlawful to use “any device, scheme, or artifice to defraud,” to make material omissions or misrepresentations, or to engage in a fraudulent act “in connection with the purchase or sale of any security.” Section 10(b) and Rule 10b-5 do not require that a security transaction in question be effected on a national exchange, but they do require proof of scienter. Negligence has not been held sufficient to establish scienter under Rule 10b-5, but courts have held that proof of recklessness is adequate to establish a cause of action.

In addition to the general securities laws that the SEC can utilize to regulate short selling, the SEC has also made permanent Rule 204T, which was initially adopted as a temporary measure in October 2008. Rule 204 now imposes a borrowing delivery requirement to try to reduce “potentially abusive ‘naked’ short selling.” This rule requires market participants who fail to deliver securities at settlements to close out the position (by borrowing or purchasing securities).

\[\text{References}\]

198 Rule 10b-5 is used so commonly because it is also the rule by which the SEC enforces violations of insider trading. See SEC v. Zandford, 535 U.S. 813, 819-20 (2002); SEC v. Tex. Gulf Sulphur, 401 F.2d 833, 847-50 (2d Cir. 1968) (en banc).
201 See Aaron v. SEC, 446 U.S. 680, 691 (1980). (“[S]cienter is an element of a violation of § 10(b) and Rule 10b-5, regardless of the identity of the plaintiff or the nature of the relief sought.”).
202 Sundstrand Corp. v. Sun Chem. Corp., 553 F.2d 1033, 1045 (7th Cir. 1977) (“[R]eckless conduct may be defined as . . . involving not merely simple, or even inexcusable, negligence, but an extreme departure from the standards of ordinary care, and which presents a danger of misleading buyers or sellers that is either known to the defendant or is so obvious that the actor must have been aware of it.” (quoting Franke v. Midwestern Okla. Dev. Auth., 428 F. Supp. 719, 725 (W.D. Okla. 1976))).
203 See July 2009 Final Rule Release — Rule 204, supra note 29, at 38,266.
204 Id.
by the start of the trading day following the settlement date.\footnote[205]{Id. at 38,292.} If the failure to deliver is not timely closed out, then the participant

may not accept a short sale order in the equity security from another person, or effect a short sale in the equity security for its own account, to the extent that the broker or dealer submits its short sales to that participant for clearance and settlement, without first borrowing the security, or entering into a bona fide arrangement to borrow the security, until the participant closes out the fail to deliver position by purchasing securities of like kind and quantity and that purchase has cleared and settled at a registered clearing agency.\footnote[206]{Id. at 38,292 (quoting 17 C.F.R. § 242.204(b)).}

This rule is particularly designed to address potentially abusive naked short selling, which can also affect investor confidence and create “unwarranted reputational damage.”\footnote[207]{Id. at 38,267-68.} The SEC clearly has a variety of enforcement provisions available to address fraud or price manipulation through short selling, but these rules all require targeted enforcement, while a price test rule theoretically provides a general backstop for abusive practices without requiring the SEC to prove intentional conduct or targeted manipulative action toward a specific security.

V. \textsc{The Importance of a Price Test Rule in Today’s Markets}

At the time the uptick rule was repealed in 2007, the SEC felt that short selling regulation could be scaled back because of the “high levels of transparency and regulatory surveillance” in modern markets.\footnote[208]{2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,069.} The Commissioners felt that the “abusive or manipulative” short selling the uptick rule was designed to curb was less likely to occur in what the agency felt to be a highly-regulated environment.\footnote[209]{Id.} However, during the eighteen months following the removal of the uptick rule, a series of events occurred, indicating that modern U.S. markets were neither transparent nor well-regulated. The federal government was required to bail out the United States
banking system due to toxic securities that threatened liquidity; one major investment bank was sold at a fire-sale price to avoid collapse, while another investment bank holding company filed for Chapter 11; and the federal government was forced to take an equity stake in the world's largest insurance company as a result of the company's liabilities for credit default swaps. The proverbial icing on the cake was the discovery of a massive Ponzi-style fraud scheme costing investors untold billions of dollars and perpetrated by a well-known financier. While these incidents were unrelated to short selling per se, and despite the many securities regulations in place, modern markets appeared to be far less transparent than posited when deciding to repeal the uptick rule.

A. Why Regulation in the Absence of a Price Test Rule Is Insufficient

The fact that the SEC relied on its assumption of transparent and well-regulated markets when it repealed the uptick rule is troublesome. The SEC performed the Pilot and reviewed several academic studies based on the Pilot before repealing the uptick rule, all of which indicated that repeal of the rule would not affect investors and the market. However, after the markets proved to be inadequately regulated in 2008, the SEC was tasked with reevaluating the uptick rule or

---


211 In December 2008, Bernard Madoff was accused of bilking investors out of billions of dollars through his investment fund. See generally Ross Kerber, The Whistleblower, BOSTON GLOBE, Jan. 8, 2009, available at http://www.boston.com/business/articles/2009/01/08/the_whistleblower/. The SEC was alerted by at least one tipster who had worked for a competing investment fund, several times over the eight-year period before the scheme was uncovered. Id. Mr. Madoff was investigated numerous times by the SEC but was never charged with a single offense. Id. This scheme was not one that affected only unsophisticated investors. Major banks, investment funds and institutions lost money in the scam, such as Banco Santander, Union Bancaire Privee, HSBC, BNP Paribas, Fairfield Greenwich, Yeshiva University and New York Law School. Madoff's Victims, WALL ST. J., Jan. 28, 2009, available at http://s.wsj.net/public/resources/documents/st_madoff_victims_20081215.html. As a result of Madoff's actions, many other investors sought to recoup their money from other investment funds and similar Ponzi schemes have been uncovered. David Glovin & Joe Schneider, Nadel, Missing Hedge Fund Adviser, Arrested by FBI, BLOOMBERG, Jan. 27, 2009, available at http://www.bloomberg.com/apps/news?pid=20601087&sid=a64niaXcacs&refer=home.

212 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,069.
coming up with another type of price test regulation to police manipulative or fraudulent short selling and to preserve investor confidence in a volatile market.

1. Importance of Investor Protection

After various corporations, hedge funds, and individual 401(k)s were wiped out in 2008, there was a call for the SEC to get back to basics by focusing on investor protection.\(^\text{213}\) Reinstating a price test rule for short selling is one way for the SEC to enhance investor protection. A major underlying goal of all federal securities laws is to protect investors by preventing price manipulation in securities markets.\(^\text{214}\) In its Proposed Rule Release discussing the results of the Pilot, the SEC noted that while “there is concern regarding the possibility of manipulation using short sales,” the Pilot report did not note any increases in this practice during the Pilot period.\(^\text{215}\) However, the Pilot report claims that the analysis it performed was not designed to directly examine whether instances of market manipulation through short selling occurred during the Pilot.\(^\text{216}\) The OEA also noted that there was a possibility that “traders with manipulative intentions” may have been more reluctant to act during the Pilot because of the additional layer of analysis of the trades due to the Pilot.

In spite of this possibility, one of the major reasons the SEC felt comfortable removing the uptick rule was the availability of the “general anti-fraud and anti-manipulation


A fundamental goal of the federal securities laws is the prevention of manipulation. Manipulation impedes the securities markets from functioning as an independent pricing mechanism, and undermines the integrity and fairness of those markets. Congress granted broad rulemaking authority to the Commission to combat manipulative abuses in whatever form they might take, including anti-fraud, prophylactic, and general rulemaking authority.

\(^{215}\) 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,069, n.16.

\(^{216}\) OEA PILOT STUDY, supra note 80, at 47-48 (“The type of analysis conducted in this study cannot directly prove whether market participants are engaging in manipulative practices, because it is inherently difficult to measure whether the Pilot has had any impact on the degree to which markets are susceptible to manipulation.”).
provisions under the Securities Act and the Exchange Act that allow the agency to enforce fraud or price manipulation effected through short selling. While the SEC can utilize these statutes and rules to prevent manipulation and fraud in the absence of a price test rule for short selling, these regulations have several elements that must be established in order for the SEC to prove that fraud or manipulation related to short selling actually occurred. Proving fraud or manipulation can be particularly problematic when delineating whether or not information is a rumor. Trading and investment decisions on Wall Street are based partially on research and partially on instinct and opinions. At times, opinions may also be nothing more than thinly-veiled rumors. The difficulty in regulating the spread of rumors is especially obvious today, when the sheer number of communication methods available allows rumors to spread like wildfire. Regulating the flow of information in the market is a mammoth task. Thus, a permanent backstop to prevent manipulative short selling would supplement the other securities regulations and give investors, companies, and markets as a whole consistent protection.

2. Whispers on the Street—the Rumor Problem

The difficulty of regulating rumors in the markets has been acknowledged by the SEC. Testifying before the Senate Banking Committee in July 2008, then-SEC Chairman Christopher Cox admitted that the SEC did not historically bring enforcement action against those spreading false rumors about a stock because of the complexities involved in determining who originated a false rumor and whether that originator knew that the information he or she spread was false. During this testimony, Mr. Cox also pointed to the Commission’s April 2008 lawsuit against a trader as a prime example of the SEC’s new attempts to enforce the spread of false rumors. In April 2008, the SEC filed a suit against Paul

217 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,075.
218 Id.
219 See supra Part IV.
221 Id.
S. Berliner, a trader, for allegedly spreading a false rumor about a pending acquisition transaction and charged him with securities fraud and market manipulation.\(^{222}\) In the case, the SEC used records of electronic communications, specifically instant messages, to obtain evidence of the fraud and manipulations.\(^{223}\) The complaint in this matter alleges that the trader spread rumors about the pending acquisition of Alliance Data Systems Corp. (“ADS”) by the Blackstone Group.\(^{224}\) ADS had agreed to sell its shares to Blackstone for $81.75 per share.\(^{225}\) The SEC alleged that the trader sent instant messages to thirty-one investment professionals claiming Blackstone amended its per-share offer and would now offer only $70.00 per share for ADS’s stock.\(^{226}\) The media eventually picked up on this rumor and further disseminated the alleged misinformation.\(^{227}\) To illustrate the incredible swiftness with which a rumor can spread through the modern financial world, the stock price of ADS dropped \textit{seventeen percent} in the \textit{thirty minutes} after the trader sent the first instant message.\(^{228}\) Later the same day, the NYSE put temporary curbs on trading in ADS stock, and ADS was forced to issue a press release denying the rumor.\(^{229}\) Concurrently with sending instant messages, the trader also began to short sell 10,000 shares of ADS stock, for a profit of $25,509.\(^{230}\)

This case was eventually settled,\(^{231}\) but is notable for several reasons. First, the availability of electronic evidence in this case made the SEC’s task of proving who originated the rumor, whether or not the rumor was material information,

\(^{222}\) Complaint at 1-2, SEC v. Berliner, 08-CV-3859 (Sec. & Exch. Comm’n Apr. 24, 2008). Mr. Berliner was charged with fraud under sections 17(a) of the Securities Act and 10(b) of the Exchanges Act (and Rule 10b-5 there under) and with market manipulation under section 9(a)(4) of the Exchange Act. Id. at 6-8. See SEC Litigation Release No. 20,537, Apr. 24, 2008, http://www.sec.gov/litigation/litreleases/2008/lr20537.htm; see also supra Part IV for more information on these statutes.

\(^{223}\) Id., supra note 222, at 4.

\(^{224}\) Id. at 1.

\(^{225}\) Id.

\(^{226}\) Id. at 4.

\(^{227}\) Id. at 5.

\(^{228}\) Id. at 4.

\(^{229}\) Id. at 4-5.

\(^{230}\) Id. at 5-6.

\(^{231}\) The trader accused agreeing to settlement without admitting or denying wrongdoing. SEC Litigation Release No. 20,537, supra note 222. He had to pay back his illicit profits, pay a civil penalty of $130,000 and he was banned from “association with any broker or dealer.” Id.
and whether or not there was scienter\footnote{See supra Part IV.} far easier to prove than if the information in question was spread with an offhand comment made during a phone call or over a business lunch. This type of concrete evidence certainly will not be available in every case of short selling fraud or manipulation, which further hampers SEC enforcement attempts. Second, the trader in this case only made a profit of $25,509.\footnote{See Kara Scannell & Gregory Zuckerman, SEC Accuses Ex-Trader of Blackstone Ruse, WALL ST. J., April 25, 2008, at C1 (“SEC staff searched data embedded in electronic communications, called meta data, which can trace emails from sender to sender.”).} While not negligible, this amount is far from remarkable. However, while making this nominal illicit profit of just $25,000, the trader caused the total market capitalization of ADS to decline by nearly $1.2 billion in just thirty minutes.\footnote{Market capitalization is the value of a company’s outstanding shares and is calculated by multiplying the share price of a security by the total number of shares outstanding. Forbes Investopedia, Market Capitalization, http://investopedia.com/terms/m/marketcapitalization.asp (last visited Jan. 22, 2009). As of December 31, 2007, in its audited financial statements, ADS had 87,786 million shares issued, which is the estimate that will be used for this calculation. Alliance Data Systems, 2007 Annual Report (Form 10-K), at F-5 (Feb. 28, 2008), available at http://www.sec.gov/Archives/edgar/data/1101215/000119312508041317/d10k.htm#toc41827_21. Using the share prices contained in the SEC’s complaint, if ADS’s stock dropped from approximately $77.00 to $63.65, this is a difference of $1.17 billion dollars—($77 \times 87,786MM = $6.76B) less ($63.65 \times 87,786MM = $5.59B)). Complaint, supra note 222, at 4.}

Reliable information is paramount to market confidence. The decline in the market capitalization of ADS stock related to the 
\textit{Berliner} case illustrates how quickly markets can react and the devastating effect certain information can have on the price of a security.\footnote{Testimony Concerning Recent Developments in U.S. Financial Markets and Regulatory Responses, supra note 220.} A price test rule to regulate short selling may help mitigate potential consequences of the spread of false information or merely unconfirmed information. Although short sellers provide a positive service by bringing securities prices close to equilibrium, harmful rumors, true or not, can wreak havoc and cause extreme volatility. ADS’s stock price recovered by the end of the trading day in question as a result of the company taking quick action to quash the rumor, but the fact remains that the information that drove down the stock price spread all over Wall Street in a matter of minutes.\footnote{Complaint, supra note 222, at 4.}

When rumors persist over a longer period of time and are not blatantly fraudulent on their face or easily rebuttable
by a company (like the rumor in the Berliner case), those rumors could potentially have a more devastating effect. One method of self-help against abusive short selling is for a company to repurchase its own stock, creating a short squeeze, and showing that the company has access to the capital markets. In fact, at least one commentator has blamed the SEC for preventing companies from protecting themselves from manipulative short sellers in this manner because of uncertainty about whether the companies themselves will be charged with illicit market manipulation. In cases of rumors spread over a protracted period, a price test on short selling could slow down the race to short a company’s stock, perhaps allowing a company a bit more time to deal with their financial issues in a more orderly fashion, ensuring compliance with all other SEC regulations and ultimately protecting investors.

In the thick of the 2008 credit crisis, politicians and Wall Street executives speculated that protracted rumors were what killed Bear Stearns. On July 15, 2008, the SEC issued a warning that it would begin investigating these rumors to ensure that there was no fraud or collusion occurring to drive down Bear Stearns’s stock price. Similar concerns regarding rumors were voiced in regard to Lehman Brothers in the months before the bank collapsed. Since Wall Street “deals in rumors,” regulators have no easy task in proving that short sellers knowingly spread false information in order to profit. The SEC has maintained throughout the credit crisis that it was investigating accusations of fraud and price manipulation through short selling, but because some enforcement

237 See supra note 35 and accompanying text.
238 Macey, The Government is Contributing to the Panic, supra note 61.
239 Id. The SEC publicly acknowledged this problem after the demise of Bear Stearns and Lehman. Id.
240 See Anderson, supra note 54.
241 See July 2008 Emergency Order, supra note 29, at 47,379. (“False rumors can lead to a loss of confidence in our markets. Such loss of confidence can lead to panic selling, which may be further exacerbated by ‘naked’ short selling . . . During the week of March 10, 2008, rumors spread about liquidity problems at Bear Stearns, which eroded investor confidence in the firm. As Bear Stearns’ stock price fell, its counterparties became concerned, and a crisis of confidence occurred late in the week. In particular, counterparties to Bear Stearns were unwilling to make secured funding available to Bear Stearns on customary terms.”). The order also effected a ban on naked short sales of 19 financial stocks that was effective immediately and ended on July 29, 2008. Id.
242 See Louise Story, Lehman Being Besieged by Investor Betting Against It, N.Y. TIMES, June 4, 2008.
243 Anderson, supra note 54.
investigations can take years, it will be some time before the findings of these investigations come to light.

B. The Uptick Rule as a Backstop

Some critics of reinstating the short sale price test rule have pointed to the SEC's recent implementation of Rule 204, regulating naked short sales, as another method to combat manipulation through short selling, rendering a price test rule unnecessary. While this rule will certainly reduce the number of “fails to deliver” and thus reduce the chances of manipulation through short selling, Rule 204 does not provide an overall backstop like a price test rule and thus would limit investor protection.

The uptick rule was also designed to prevent short selling from hastening a decreasing market so that even down markets would remain orderly. However, during the twelve month span of the Pilot, the Standard & Poor's 500 Index increased nearly 12.8%, while the Dow Jones Industrial Average was up 10.9%. The Pilot study occurred when the market was increasing in an “orderly” fashion, but provided no data that the SEC could use to make an informed conclusion about the effect of the repeal of the uptick rule in volatile, declining markets, similar to those of the third and fourth quarters of 2008.

The sharp increase in short selling that occurred at the height of the credit crisis in 2008 was exactly the scenario the NYSE warned of in its 2006 comment letter, even though it ultimately supported repeal of the rule.

---

244 See supra notes 93-96 and accompanying text. The defendant was accused of illicit short selling in August 2005, and the complaint was not filed until October 2007.
245 See, e.g., Letter from Paul M. Russo to Elizabeth Murphy, supra note 158.
246 See supra notes 205-207 and accompanying text.
247 See, e.g., 2006 Proposed Amendments to Regulation SHO and Rule 10a-1, supra note 46, at 75,070.
250 Opposing Uptick Rule Is Truly Short-Sighted, supra note 132.
251 Letter from Mary Yeager, Assistant Sec'y, N.Y. Stock Exch., to Nancy M. Morris, Sec'y, SEC, supra note 149.
author of an independent empirical study using Pilot test data warned that the Pilot results only supported the SEC’s decision to repeal the uptick rule “conditionally, . . . the condition being the absence of extreme market conditions.”

It would be simplistic to blame the increase in short selling and market volatility following the repeal of the uptick rule on the removal of the rule alone; many factors contributed to the upheaval, creating a “perfect storm” of sorts. Some have said that they “doubt whether the continued existence of the short sale rule is justified, other than perhaps to provide investors with a semblance of confidence in the markets.” However, the absence of the uptick rule or any other price test restriction or circuit breaker on short selling during a period of high market volatility and declining stock prices was seen by some as a factor in exacerbating this credit crisis. Even if the uptick rule is viewed as a mere prophylactic measure, allowing investors to retain some “semblance of confidence” in a falling market, this underlying confidence can be helpful in calming volatile markets and stabilizing cascading stock prices.

The volatility and uncertainty in the market is what led the SEC to issue its September 2008 short sale ban. The temporary ban on short sales of financial stocks was thrown in place to prevent the collapse of more financial institutions and further liquidity issues. Despite the lip service paid to short sellers by the SEC about the positive benefits of short sales, such as preventing bubbles, properly valuing stock, and detecting corporate fraud, short sellers were among the first blamed during the financial crisis. One major complaint after the SEC announced this ban was whether financial stocks could be properly valued. For example, on September 15, 2008, the stock price of Washington Mutual closed at $2.00. When the short sell ban took effect on September 19, the stock price of the company, which was included on the banned list, closed at $4.25. On September 25, the share close price was $1.69.

---

254 Macey et al., *supra* note 27, at 805.
255 Schwab, *supra* note 213.
256 Macey et al., *supra* note 27, at 805.
258 Bloomberg Investment Service, Historical Quotes.
259 *Id.*
and that evening, the federal government seized the bank and sold its assets to JPMorgan Chase. From the date the short sell ban took effect until the day the bank failed, its stock price never dipped below $2.26 per share, 13% above its value on September 15, the day that Lehman declared Chapter 11. On September 16, Washington Mutual issued a statement saying that it “shouldn’t be judged by its stock price,” and that the bank would not “go the way of Lehman Brothers” even though there were already rumors circulating that JPMorgan Chase was considering a merger.

In this case, the shorts were right about Washington Mutual, and it is conceivable that the short sale ban artificially prolonged the company’s life, but in light of the crisis surrounding the markets, this was not necessarily a negative result. Any breathing room provided by the short sale ban may have allowed the relatively orderly demise of Washington Mutual. In a market regulated by a short sale price test rule, the hope is that an extreme full-out ban on short sales would not be necessary to achieve this result. Ideally, a short sale price test rule would provide a backstop during troubled times, without impinging on beneficial short selling in an advancing market.

VI. CONCLUSION

After a period of market deregulation beginning in the mid-1990s, the financial crisis of 2008 devastated the American economy and resulted in an increase in support for more regulation of financial markets. The credit crisis of 2008 was certainly not caused by short sellers; bear raids and stock price manipulation were not what put the global economy on the precipice of collapse and the existence of the uptick rule during this period certainly would not have been a panacea to heal all market woes. Investors continued to lose confidence once the

260 Id.
261 Id.
global securities markets started cascading downward. In any
case, the uptick rule or other price test backstop could have
certainly helped instill some confidence in the stability of the
markets. The fact that the uptick rule remained intact and
unchanged since the advent of modern securities regulation
until 2007 is remarkable, particularly considering that it was
challenged and debated in so many instances.\textsuperscript{265} Perhaps a price
test rule for short selling is merely a placebo to make investors
feel better during tough times, but that does not mean the rule
cannot be effective; it is difficult to argue that confidence in the
securities markets is unimportant.

The SEC’s dire decision to place an emergency ban on
the short sale of securities of nearly 1000 companies in
September 2008 points to the fact that either short selling was
cauing problems in a declining, nervous market environment
or that there was a perception among market participants that
this was the case. The decision to ban short selling of certain
stocks is the sort of panicky and rash decision-making in a
decaying market that the uptick rule was designed to
prevent.\textsuperscript{266} Depending on regulators to piece together a last-
minute ban in emergency situations in lieu of a permanent
price test is an unwise decision. The SEC has powers to pass
emergency orders when it sees fit, but waiting until a crisis
occurs before taking regulatory action is not a wise practice,
particularly if the action is too little or too late. Further,
defining an emergency is not always simple.

Finally, relying solely on fraud or anti-manipulation
provisions of the Securities and Exchange Acts to combat
abusive short selling will be a losing proposition for the SEC,
even with the addition of a rule to prevent naked short selling.
The agency is already strapped for resources,\textsuperscript{267} but more salient
is the fact that stock markets deal regularly with rumors and
opinions. It is nearly impossible to stop whispers on the Street.
The SEC has gone so far as touting its actions concerning the
\textit{Berliner} case on its website as a “landmark” action that shows
the Commission was “[a]gressively [c]ombatting” market

\textsuperscript{265} See supra notes 107-110 and accompanying text.
\textsuperscript{266} See supra note 90 and accompanying text.
\textsuperscript{267} See Robert Chew, \textit{Can Mary Schapiro Revitalize the SEC?}, TIME.COM, Jan.
SEC’s declining levels of enforcement staff).
manipulation during the credit crisis. If this one case, which ended in settlement, without adjudication, is a “landmark,” it seems that using enforcement actions as the only way to prevent market manipulation through short selling will likely prove to be very difficult and ultimately, ineffective.

After considering several iterations of a short sale price test rule, the SEC settled on Rule 201, an alternative version of the original uptick rule to help protect investors and restore confidence. The implementation of Rule 201 should help prevent short sale manipulation, but some feel the rule does not go far enough. Moreover, the Rule is likely to be challenged by those who believe that regulation of short selling is market-restricting and comes at too high of an economic cost. It will thus remain to be seen if the new rule both proves effective in combating abusive short selling practices and achieves the same longevity as the original uptick rule.

Melissa W. Palombo

---

7 J.D., Brooklyn Law School, 2010; B.S. Accounting and Business Administration, Washington and Lee University. Many thanks to my colleagues at the Brooklyn Law Review, especially Andrei Takhteyev, Joseph Roy, Joseph Lanzkron, and Jonathan Perrelle who all provided invaluable feedback. Special thanks to my fiancé Brannon Cook and my wonderful family for their support and encouragement.
A Right of Confrontation for Competition Hearings Before the European Commission

Procedural fairness lies at the bedrock of justice.¹ Among all procedural rights in the American judicial system, the right of a criminal defendant to confront and cross-examine government witnesses is one of the most vital.² Cross-examination allows criminal defendants to test the perception and memory of witnesses, while giving a neutral fact-finder a first-hand view of the witnesses’ consistency, credibility, and biases.³ As Wigmore noted, it is “the greatest legal engine ever invented for the discovery of truth.”⁴

On September 12, 2009, Christine Varney, the United States Assistant Attorney General for Antitrust, called on “competition agencies, international organizations, and the antitrust community to discuss procedural fairness more broadly, focusing on the opportunity to refine procedures that parties can understand and rely on as a means of removing unnecessary uncertainty from enforcement efforts.”⁵ Nowhere is procedural fairness more important in competition law than in cartel enforcement, an area of the law dealing with agreements among competitors to restrain trade through actions such as “price-fixing, market allocation, and bid-rigging,”⁶ where fines and prison sentences can have catastrophic consequences for defendants.⁷

³ See Fred O. Smith, Crawford’s Aftershock: Aligning the Regulation of Nontestimonial Hearsay with the History and Purposes of the Confrontation Clause, 60 STAN. L. REV. 1497, 1518 (2008).
⁴ 3 JOHN HENRY WIGMORE, A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW § 1367 (2d ed. 1923).
⁵ See Varney, supra note 1, at 4.¹
⁶ See ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT, FIGHTING HARD CORE CARTELS: HARM, EFFECTIVE SANCTIONS AND LENIENCY
⁷
This note focuses on the adequacy of procedural safeguards in cartel enforcement proceedings in the European Union ("EU"). Specifically, it will investigate the European Commission's ("EC" or "Commission") use of "paperless" applications for leniency from fines and oral statements gathered during cartel investigations absent a right to confront and cross-examine witnesses. Because co-conspirators are rewarded with leniency based upon either the sufficiency or "added value" of the information they offer to the Commission, they have an incentive to paint this evidence in a light least favorable to their fellow cartel members and most favorable to themselves. Without a mechanism for targeted corporations to directly ascertain the truth of those allegations, the use of oral statements as evidence in EC cartel enforcement proceedings raises a number of the classic evils against which the rights of confrontation and cross-examination are designed to protect. Recently, commentators have argued that because oral statements in the EC are unsworn, declarants are not subject to compulsory process, and the prosecuted parties are not afforded an opportunity to confront and cross-examine declarants, the use of these statements as evidence violates the rights that targets of EC cartel enforcement proceedings should enjoy. However, the solutions that these critics have offered go too far—requiring a wholesale reform of EC competition


10 See discussion infra Part II.B.1.

11 Forrester, supra note 7, at 833.

12 See discussion infra Part IV.A.

13 See discussion infra Part IV.A.

14 See discussion infra Part IV.B.
procedure. This note proposes a more conservative solution grounded in the right of confrontation jurisprudence of the United States.

Part I of this note will provide a brief historical overview of the global development of competition law, cartel enforcement, and leniency programs, and will present the United States’ approach as a model. Part II will outline the relevant competition laws of the EU, as well as the standards of procedural fairness that currently undergird those laws. Part III will argue that the current procedural safeguards have not kept pace with the EC’s substantive and procedural trends toward a criminal and adversarial system. Lastly, Part IV will survey competing proposals for procedural reforms in the EC and will offer a counterproposal. Specifically, this note will argue that the correct solution to the inadequacy of the current EC competition hearing procedures is to amend the regulations governing cartel enforcement to allow for a right to confront and cross-examine witnesses modeled on the American approach, as outlined in the Supreme Court’s decision in Crawford v. Washington. 15

I. BACKGROUND AND THE AMERICAN MODEL

In order to understand the current European cartel enforcement regime, it is first important to understand the goals and development of competition laws. In general, competition laws are designed to promote competition in market economies for the benefit of consumers, 16 a goal achieved through two basic means: (1) laws prohibiting excessive market power, and (2) laws designed to deter and prosecute unfair competition among competitors. 17 The first category includes the regulation of mergers and acquisitions, as well as anti-monopoly laws, while the second category, which is the subject of this note, concerns the preservation of competition and market fairness. 18 Additionally, a full understanding of modern European cartel enforcement requires a familiarity with the cartel enforcement regime of the

16 See Eleanor M. Fox, Linked-In: Antitrust and the Virtues of a Virtual Network, 43 INT’L LAW. 151, 152 (2009) [hereinafter Fox, Linked-In].
18 Id.
United States, which remains a dominant influence in the shaping of EC competition policies and procedures.19

A. Background

In response to the proliferation of industrial combinations, known as “trusts,” the United States enacted the Sherman Act of 1890, which was designed to (1) “prevent[] the high prices associated with monopoly or cartel activity” and to (2) protect[] the right of every person to practice a trade of choice.”20 For the next fifty years, aggressive competition enforcement was largely limited to the United States.21 However, the second half of the twentieth century saw a significant increase in competition laws globally.22 Today, competition laws are enforced in over 100 jurisdictions,23 including multiple cross-border competition “networks.”24 One area of competition law, cartel enforcement, is executed through a number of investigative and deterrent mechanisms designed to ferret out cartel activity, which is notoriously secretive.25 And while global competition authorities differ in their approaches to cartel enforcement, one mechanism has proved especially effective at cartel detection and enforcement: the “leniency program.”26

Under a leniency program, a competition authority encourages cartel members to self-report anticompetitive

21 Fox, Linked-In, supra note 16, at 152.
22 See id.
23 Id. at 154.
24 Fox, Antitrust and Regulatory Federalism, supra note 17, at 1782.
26 O.E.C.D., FIGHTING HARD CORE CARTELS supra note 6, at 7. (“Leniency programs uncover conspiracies that would otherwise go undetected and also make the ensuing investigations more efficient and effective. Experience shows that these programs work.”); Julian M. Joshua, That Uncertain Feeling: The Commission's 2002 Leniency Notice, EUROPEAN COMPETITION LAW ANNUAL 2006: ENFORCEMENT OF PROHIBITION OF CARTELS 512 (Claus-Dieter Ehlermann & Isabela Anatnasiu eds., 2006) (“[N]o self-respecting antitrust agency with any aspiration to effective enforcement is without a leniency policy.”).
conduct in return for a conditional promise either to refrain from bringing criminal charges or to reduce potential fines.\textsuperscript{27} Leniency programs destabilize cartels by creating a “prisoner’s dilemma,” fostering distrust among co-conspirators and providing an incentive for each member to turn in its fellow cartel participants to competition authorities.\textsuperscript{28} The leniency program has been described by American officials as “[u]nquestionably . . . the greatest investigative tool ever designed to fight cartels.”\textsuperscript{29} The first cartel leniency program was developed in 1978 in the United States.\textsuperscript{30} Today, nearly fifty countries have enacted leniency programs.\textsuperscript{31} These programs have been highly successful, resulting in extraordinary fines and significant prison sentences.\textsuperscript{32}

\textbf{B. The American Model}

The United States takes an aggressive approach to criminal cartel enforcement. The Assistant Attorney General for the Antitrust Division of the Department of Justice has stated that cartel conduct is “unambiguously harmful.”\textsuperscript{33} The United States Supreme Court has described cartels as “the supreme evil of antitrust.”\textsuperscript{34} Government officials have justified criminal prosecution of cartel offenses on a number of grounds, including the secretive nature of cartels,\textsuperscript{35} the existence of


\textsuperscript{28} See Roger W. Fones, Rony P. Gerrits & Nicole D. Devero, \textit{Antitrust Leniency Programs and their Impact on the Aviation Industry}, AIR & SPACE L., at 1, 19 (2008) (internal quotations omitted); Leslie, \textit{supra} note 27 at 455.


\textsuperscript{31} See Chavez, \textit{supra} note 30, at 853.

\textsuperscript{32} See, \textit{e.g.}, id. at 847-48.


\textsuperscript{34} Verizon Commc'n, Inc. v. Law Offices of Curtis V. Trinko, LLP, 540 U.S. 398, 408 (2004).

\textsuperscript{35} \textit{International Cartels Roundtable, supra} note 19, at 96. James Griffin, who was then the Deputy Assistant Attorney General for the Antitrust Division of the U.S.
criminal intent when participating in a cartel, and the inherently anticompetitive conduct of cartels.\textsuperscript{36} The American impulse to view cartel conduct as criminal has “been part of the American antitrust system from the beginning,”\textsuperscript{37} and it is an impulse that is reflected in its laws, severe penalties, and aggressive enforcement techniques.\textsuperscript{38} However, the United States also maintains a robust set of procedural protections, which guard against prosecutorial abuse and ensure respect for the fundamental rights of defendants at trial.\textsuperscript{39}

1. The Antitrust Laws

Section 1 of the Sherman Act states that “[e]very contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations, is declared to be illegal.”\textsuperscript{40} Despite the breadth of Section 1 prohibitions, U.S. criminal cartel enforcement primarily focuses on price-fixing,\textsuperscript{41} bid-rigging,\textsuperscript{42} and market allocation\textsuperscript{43} between competitors—known as the “hard-core” antitrust offenses.\textsuperscript{44} The penalties for

Department of Justice, noted that international cartels, “like all other cartels, are secret in nature. Often victims don’t even know they are being victimized, and increasingly cartel members are taking actions to ensure that they do not leave significant evidence around. . . .” \textit{Id.}

\textsuperscript{36} See Werden, supra note 25, at 5.


\textsuperscript{38} See Wils, supra note 19, at 122-24.

\textsuperscript{39} Id.


\textsuperscript{41} See U.S. DEP’T OF JUSTICE, \textit{PRICE FIXING, BID RIGGING, AND MARKET ALLOCATION SCHEMES: WHAT THEY ARE AND WHAT TO LOOK FOR} 2 (2005), available at http://www.justice.gov/atr/public/guidelines/211578.pdf (“Price fixing is an agreement among competitors to raise, fix, or otherwise maintain the price at which their goods or services are sold. It is not necessary that the competitors agree to charge exactly the same price, or that every competitor in a given industry join the conspiracy.”) [hereinafter PRICE FIXING, BID RIGGING, AND MARKET ALLOCATION SCHEMES].

\textsuperscript{42} See id. (“Bid rigging is the way that conspiring competitors [sic] effectively raise prices where purchasers — often federal, state, or local governments — acquire goods or services by soliciting competing bids.”).

\textsuperscript{43} See id. at 3 (“Market division or allocation schemes are agreements in which competitors divide markets among themselves. In such schemes, competing firms allocate specific customers or types of customers, products, or territories among themselves.”).

\textsuperscript{44} ABA \textit{SECTION OF ANTITRUST LAW, CRIMINAL ANTITRUST LITIGATION HANDBOOK} 3 (2d ed. 2006); Baker, supra note 37, at 694-95 (“Over time . . . the DOJ, and more importantly the courts, gradually developed distinguishing lines between the kinds of anticompetitive conduct that should be punished criminally and the remaining conduct, which would only be subject to civil injunctions by the government and private
criminal antitrust violations in the United States are potentially severe. Under the Antitrust Criminal Penalty Enhancement and Reform Act of 2004 ("ACPERA"), guilty corporations face fines as high as $100 million, while individuals may face up to $1 million in fines and ten years in jail. These aggressive and ever-increasing penalties have led to the collection of over $5.3 billion in fines over the last fifteen years.46

2. Powers of Investigation and the Division’s Leniency Program

The United States Department of Justice’s Antitrust Division (the “Division”) enforces the Sherman Act by conducting investigations and prosecuting offending corporations and individuals in federal court.47 Division investigations employ a wide array of aggressive techniques, including wiretaps, informants, and search warrants, to discover cartels and to build cases.48 However, above all other techniques, the Corporate and Individual Leniency Programs have been the “most effective generator of international cartel cases” for the Division.49

The Division first adopted its Corporate Leniency Program in 1978, and subsequently revised it in 1993.50

---


46 See Chavez, supra note 30, at 836.

47 Price Fixing, Bid Rigging, and Market Allocation Schemes, supra note 41, at 1.

48 Chavez, supra note 30, at 824-26.


50 Hammond, Cornerstones of an Effective Leniency Program, supra note 29, at 3 n.1. Describing the changes from 1978 to 1993, Hammond noted that “[t]he Amnesty Program was revised in three major respects. First, the policy was changed to ensure that amnesty is automatic if there is no pre-existing investigation. . . . Second,
revised program sets forth two means of achieving leniency, “Part A Leniency,” and “Part B Leniency,” which are differentiated based on whether the Division is aware of the illegal activity being reported at the time when the corporation

the Division created an alternative amnesty, whereby amnesty is available even if cooperation begins after an investigation is underway. Third, if a corporation qualifies for automatic amnesty, then all directors, officers, and employees who come forward with the corporation and agree to cooperate also receive automatic amnesty. Under Part A Leniency, the Division will grant a corporation leniency if, before an investigation has begun, it meets six criteria:

1. At the time the corporation comes forward to report the illegal activity, the Division has not received information about the illegal activity being reported from any other source;
2. The corporation, upon its discovery of the illegal activity being reported, took prompt and effective action to terminate its part in the activity;
3. The corporation reports the wrongdoing with candor and completeness and provides full, continuing and complete cooperation to the Division throughout the investigation;
4. The confession of wrongdoing is truly a corporate act, as opposed to isolated confessions of individual executives or officials;
5. Where possible, the corporation makes restitution to injured parties; and
6. The corporation did not coerce another party to participate in the illegal activity and clearly was not the leader in, or originator of, the activity.

Under Part B Leniency, if a corporation does not meet the requirements for Part A Leniency, it may qualify for leniency if—before or after an investigation has begun—it meets the following seven criteria:

1. The corporation is the first one to come forward and qualify for leniency with respect to the illegal activity being reported;
2. The Division, at the time the corporation comes in, does not yet have evidence against the company that is likely to result in a sustainable conviction;
3. The corporation, upon its discovery of the illegal activity being reported, took prompt and effective action to terminate its part in the activity;
4. The corporation reports the wrongdoing with candor and completeness and provides full, continuing and complete cooperation that advances the Division in its investigation;
5. The confession of wrongdoing is truly a corporate act, as opposed to isolated confessions of individual executives or officials;
6. Where possible, the corporation makes restitution to injured parties; and
7. The Division determines that granting leniency would not be unfair to others, considering the nature of the illegal activity, the confessing corporation’s role in it, and when the corporation comes forward.

Id. at 2-3.
first applies for leniency. In 1994, the Division also released an Individual Leniency Policy, which encourages “individuals who approach the Division on their own behalf, not as part of a corporate proffer or confession, to seek leniency for reporting illegal antitrust activity of which the Division has not previously been made aware.” Additionally, only the first corporation or individual to approach the Division is granted leniency, creating a race between conspirators.

Leniency applications to the Division may be made in writing or orally, although full cooperation inevitably requires witness statements and the production of relevant documents. The availability of oral—or “paperless”—leniency is vital to the success of the program, because “amnesty does not protect recipients from liability in private damage actions, [such that] companies would be loathe to participate or cooperate if there were a substantial risk that the evidence they provide the Justice Department could be used against them in the civil suits that inevitably follow.”

In the event that a corporation or individual is granted leniency, the Department of Justice issues a Conditional Leniency Letter, which contains the conditions of leniency and the cooperation required of that corporation or individual.

---

53 U.S. DEP'T OF JUSTICE, ANTITRUST DIV., LERNENCY POLICY FOR INDIVIDUALS 1 (1994), available at http://www.justice.gov/atr/public/guidelines/0092.pdf. In order to qualify for individual leniency, the applicant in question (1) must approach the Division before it has received any information about the anticompetitive conduct and prior to the initiation of any investigation or corporate application for leniency; (2) must report the illegal activity “with candor and completeness” and “provide[] full, continuing and complete cooperation to the Division throughout the investigation;” and (3) the individual must not have “coerce[d] another party to participate in the illegal activity” and must not have been the “leader in, or originator of, the activity.” Id. at 1-2.

54 See CORPORATE LERNENCY POLICY, supra note 51.


56 See Lazewitz & Miller, supra note 44, at 264.

57 Kolasky, supra note 55, at 213.

58 This cooperation includes seven elements, requiring the corporation to: (a) provide “a full exposition” of all the facts relevant to its anticompetitive conduct; (b) produce the remaining, non-privileged, relevant documents related to that conduct; (c) to secure and encourage current and former “directors, officers, and employees” to provide all relevant information related to the conduct; (d) to facilitate the appearance of such employees for interviews and testimony related to the reported anticompetitive activity; (e) to use its “best efforts” to ensure “complete[], candid[], and truthfull[]” responses to all interviews, grand jury appearances, and trials; (f) to use its “best efforts” to ensure that such individuals “make no attempt either falsely to protect or falsely to implicate any person or entity;” and (g) to make all reasonable efforts to pay restitution to those injured by any anticompetitive conduct reported that effects the United States. U.S. Dep't of Justice, Antitrust Division, Model Corporate Conditional
Furthermore, the Corporate Conditional Leniency Letter extends leniency to those directors, officers, and employees who provide relevant information and admit to anticompetitive activity, subject to further conditions. 59 If the applying corporation or individual meets the conditions for leniency and the Division receives the benefit of the information that the leniency applicant provides, the Division will grant Final Leniency. 60

Where the Division determines that there is sufficient evidence to proceed with a prosecution—generally where there is direct evidence of collusion—it may convene a grand jury 61 to determine whether there is enough evidence to bring formal charges against the alleged conspirators. 62 Once formal charges are brought, federal cartel cases are tried in open court before a judge and jury, with attendant procedural rights for defendants and a requirement that the alleged conduct be proven beyond a reasonable doubt. 63

3. Procedural Fairness in the United States

While the United States has historically approached cartel enforcement with vigor—whether through persistent investigation, innovative programs, or devastating penalties—its zealosity has been tempered by a robust set of procedural safeguards, including (1) constitutional safeguards, (2) statutory safeguards, and (3) policy safeguards. 64 One vital constitutional provision is the criminal defendant’s right to

59 See CORPORATE LENIENCY POLICY, supra note 51, at 4 (“If a corporation qualifies for leniency under Part A, above, all directors, officers, and employees of the corporation who admit their involvement in the illegal antitrust activity as part of the corporate confession will receive leniency, in the form of not being charged criminally for the illegal activity, if they admit their wrongdoing with candor and completeness and continue to assist the Division throughout the investigation.”).


62 Id.

63 These procedures are required under the Sixth Amendment of the United States Constitution. See U.S. CONST. amend. VI; see also Wils, supra note 19, at 124.

64 Wils, supra note 19, at 124.
confront and cross-examine the prosecution’s witnesses at trial. This right is buttressed by criminal and procedural sanctions for untruthful witness practices. One recent and notable case, United States v. Stolt-Nielsen S.A., demonstrates the necessity of these procedural safeguards.

a. Constitutional Safeguards—The Right of Confrontation

The Constitution of the United States enshrines a panoply of rights for criminal defendants at trial. These rights form the procedural backbone of the adversarial system, one which presumes the innocence of defendants and includes both a jury and “a judge who does not . . . conduct the factual and legal investigation himself, but instead decides on the basis of facts and arguments pro and con adduced by the parties.” One of the key facets of the adversarial system in the United States is the right of a defendant to confront and cross-examine witnesses against him or her. The Supreme Court has held this right to exist in the Fifth, Sixth, and Fourteenth Amendments to the Constitution.

First, the Sixth Amendment Confrontation Clause provides that “[i]n all criminal prosecutions, the accused shall enjoy the right . . . to be confronted with the witnesses against him . . . .” Centuries of Supreme Court case law have interpreted this clause as requiring “an opportunity for cross-examination by defense counsel in front of the jury, ordinarily with the defendant and the witness both in the courtroom.” The Confrontation Clause serves a number of purposes, reflecting a cold view of the inquisitorial practices of

---

65 U.S. CONST. amend. VI.
68 See U.S. CONST. amends. V, VI. These rights include the right against self-incrimination, the right to a speedy and public trial, the right to an impartial jury, and the right of the accused to confront adverse witnesses. Id.
71 U.S. CONST. amend. VI.
Continental Europe (and their brief use in England), and effectuating important practical objectives. First, the Confrontation Clause, together with its attendant right to cross-examination, serves as a shield against inquisitorial practices that were commonplace in medieval and renaissance England. These practices were first introduced through the English Marian bail and committal statutes, under which justices of the peace, acting both as investigators and prosecutors, were required to interrogate suspects and witnesses ex parte in order to determine whether to discharge or commit the suspects until trial. The results of these examinations were certified in court, and while they were not originally intended to serve as evidence against the defendant at trial, over time that practice began to invade the adversarial system.

The danger of these practices came into clear view during the trial of Sir Walter Raleigh in 1603. There, the defendant Raleigh was tried for treason based on the ex parte, signed confession of his alleged accomplice, Lord Cobham, who had been imprisoned. Because Cobham later retracted his

---

73 Crawford, 541 U.S. at 51.
74 Id. at 44-45; Margaret A. Berger, The Deconstitutionalization of the Confrontation Clause: A Proposal for a Prosecutorial Restraint Model, 76 MINN. L. REV. 557, 569-70 (1992) (“With the ascent of the Tudors, the English crown began to exercise more control over its enemies by importing techniques from the civil law in to the indigenous, essentially accusatorial, system of criminal procedure. Criminal proceedings took on a more inquisitor ial slant with the use of preliminary examinations and increased reliance on prerogative courts . . . During the sixteenth and seventeenth centuries, justices of the peace conducted preliminary examinations in ordinary criminal proceedings at common law. The justices—government officials who exercised police, administrative and judicial functions—privately interrogated the suspect, his accusers, and the witnesses against him. These examinations were then introduced into evidence to the detriment of the defendant who had neither the assistance of counsel nor the ability to call witnesses on his behalf . . . In cases of great political importance, however, the Privy Council, or the judicial members of the Council, examined the suspect and the other witnesses. At trial, proof usually consisted of reading statements that had been made out of court, such as depositions, confessions of accomplices, and letters. In his history of the common law, Stephen concluded that this prosecution on the basis of written statements ‘occasioned frequent demands by the prisoner to have his “accusers,” i.e. the witnesses against him, brought before him face to face.”’) (internal citations omitted).
75 Crawford, 541 U.S. at 44-45.
76 Id. However, some commentators have challenged the Court’s historical accuracy on this point. See, e.g., Thomas Y. Davies, What Did the Framers Know, and When Did They Know It? Fictional Originalism in Crawford v. Washington, 71 BROOK. L. REV. 105, 107-08 (2005); Randolph N. Jonakait, The Too-Easy Historical Assumptions of Crawford v. Washington, 71 BROOK. L. REV. 219, 224 (2005).
77 Crawford, 541 U.S. at 44.
78 See generally 1 CRIMINAL TRIALS 389-520 (David Jardine ed., 1850).
confession, Raleigh argued that the confession should not be used against him, and demanded to have Cobham brought before the court, believing that Cobham would not corroborate his previous confession in open court with Raleigh present. The Court refused this request and sentenced Raleigh to death. The public viewed this refusal as an egregious perversion of common law procedure, and it ultimately spurred the institution of the protective procedural rights of confrontation and cross-examination both in England and the early United States.

However, while some commentators have asserted that the purpose of the Confrontation Clause in American jurisprudence ends with “anti-inquisitorialism,” in fact, the clause finds solid grounding in practical objectives well-served by the rights. These purposes include the ability to test the perception and memory of a witness; the ability of the trier of fact to view a witness’ “demeanor and language” when subjected to questioning; the ability to focus a witness’ testimony on key issues or discrepancies; the ability to immediately challenge a witness’ story; and the ability to reveal potential biases in a witness’ account. At its core, the Confrontation Clause is “designed to promote the truth.”

---

79 At the trial, Raleigh stated, “It is now clear that he hath since retracted; therefore since his accusation is recalled by himself, let him now by word of mouth convict or condemn me.” *Id.* at 434.
80 *Crawford*, 541 U.S. at 44.
81 *Id.* at 44, 47-48; see also Jules Epstein, *Cross-Examination: Seemingly Ubiquitous, Purportedly Omnipotent, and “At Risk”*, 14 WIDENER L. REV. 427, 429-30 (2009) (“Langbein found cross-examination to be a necessary . . . response to three occurrences in the English trial system: the growing use of lawyers to present prosecutions in both the investigative and trial stages; the reward system that offered bounties to those who provided testimony establishing that a crime reached the severity (or degree of financial loss) to qualify as a felony and thus invited fraudulent testimony, the corrupt motive of which required cross-examination as an antidote; and ‘the crown witness system for obtaining accomplice evidence in gang crimes, a prosecutorial technique that created further risks of perjured testimony.’”) (citing JOHN H. LANGBEIN, THE ORIGINS OF ADVERSARY CRIMINAL TRIAL 246 (2003)).
82 Akhil Reed Amar, *Sixth Amendment First Principles*, 84 GEO. L.J. 641, 694 (1996) (describing the history behind the Confrontation Clause as one “born of revulsion against trial by affidavit.”); see also Sklansky, *supra* note 72, at 1688-94.
84 *Davis*, 415 U.S. at 316; Smith, *supra* note 83, at 1518.
85 Amar, *supra* note 82, at 649, 688.
Recently, the Supreme Court clarified the meaning of the Confrontation Clause in a series of groundbreaking cases, beginning with *Crawford v. Washington*. In *Crawford*, the Court relied on American and English common law to reach an understanding of the Confrontation Clause intended to reflect its original meaning. Specifically, the Court held that "testimonial" hearsay statements, when made by a witness who is not present at trial, are inadmissible as evidence unless (1) the witness is unavailable and (2) the defendant previously had the opportunity to cross-examine the witness. The Court broadly defined a "witness" as any person who "bears testimony," and further defined testimony as "typically '[a] solemn declaration or affirmation made for the purpose of establishing or proving some fact.'" As a result, the *Crawford* standard rejected the test set forth in an earlier case, *Ohio v. Roberts*, which admitted hearsay statements of witnesses not available for cross-examination so long as those statements bore sufficient "indicia of reliability." In overruling *Roberts*,
the Crawford Court argued that “[r]eliability is an amorphous, if not entirely subjective, concept,” such that the Roberts test is “permanently” unpredictable and does not adequately conform to the purposes of the Confrontation Clause. Thus, despite the criticism of the Crawford majority’s originalist interpretation of the Confrontation Clause, commentators have praised the majority’s “more principled” approach.

The applicability and necessity of cross-examination have also been recognized in the context of civil and administrative hearings, grounded in the Fifth and Fourteenth Amendment Due Process Clauses. In Goldberg v. Kelly, the Supreme Court held that the Due Process Clause of the Fourteenth Amendment required “the opportunity to be heard,” that is, a hearing “at a meaningful time and in a meaningful manner.” Additionally, the Court noted that “[i]n almost every setting where important decisions turn on questions of fact, due process requires an opportunity to confront and cross-examine adverse witnesses.” Later, in Mathews v. Eldridge, the Court further clarified the Goldberg standard in the context of the Fifth Amendment, creating a test under which courts balance three factors to determine whether procedural safeguards in a given case are sufficient; namely, (1) “the

Cobham put before a grand jury would differ significantly in the content from the accusation he made against Sir Walter Raleigh. The prosecutor has an incentive to lean on the prospective witness to shape the grand jury testimony in accordance with the prosecution’s theory of the case in order to secure an indictment and to freeze the witness’s story as much as possible.”).
private interest that will be affected by official action;” (2) “the risk of an erroneous deprivation of such interest through the procedures used, and the probable value, if any, of additional substitute procedural safeguards;” and (3) “the Government’s interest, including the function involved and the fiscal and administrative burdens that the additional or substitute procedural requirement would entail.” As such, whether or not a strict, incarceratory liberty interest is involved, the American system constitutionally recognizes the need for confrontation and cross-examination in judicial proceedings.

b. Statutory and Policy-based Protections

Defendants’ rights are also protected through statutory and contractual provisions designed to ensure witness truth-telling. First, the Federal Rules of Evidence (“Federal Rules”) provide a number of protections for defendants in criminal and civil cases, including (1) a requirement that all witnesses must swear or affirm to “testify truthfully,” and (2) rules bearing on hearsay testimony, which bar out-of-court statements offered for the truth of those statements absent one of the enumerated exceptions to or exemptions from the Federal Rules. Furthermore, non-evidentiary statutory law prohibits perjury, the making of false statements or declarations, actions in contempt of court, and obstruction of justice. Lastly, in the cartel enforcement context, the Division’s leniency agreements include contractual requirements of witness appearance and truthfulness as conditions of leniency. The practical effect of these statutory rules and contractual provisions in the context

---

101 FED. R. OF EVID. 603 (“Before testifying, every witness shall be required to declare that the witness will testify truthfully, by oath or affirmation administered in a form calculated to awaken the witness’ conscience and impress the witness’ mind with the duty to do so.”).
102 FED. R. EVID. 801-07.
104 For example, the Division’s Model Conditional Leniency Agreements note that witnesses:

when called upon to do so by the United States, [testify] in trial and grand jury or other proceedings in the United States, fully, truthfully, and under oath, subject to the penalties of perjury (18 U.S.C. § 1621), making false statements or declarations in grand jury or court proceedings (18 U.S.C. § 1623), contempt (18 U.S.C. §§ 401-402), and obstruction of justice (18 U.S.C. § 1503-1521), in connection with the anticompetitive activity being reported.
of the Division’s leniency program is to strongly discourage leniency applicants from providing false or misleading testimony, to penalize those corporations and individuals who seek to obtain leniency or punish co-conspirators through dishonesty, and to ensure compliance with the above-mentioned constitutional provisions, including the Confrontation Clause. In fact, the Division has used allegedly untruthful statements as a basis for revoking conditional leniency, as in the case of United States v. Stolt-Nielsen S.A. However, the true lesson of Stolt-Nielsen lies in its revelation of both the pervasiveness of witness dishonesty and the necessity of a right of confrontation in cartel proceedings.

c. The Case of United States v. Stolt-Nielsen S.A.

The necessity for the right of a defendant to confront adverse witnesses has been demonstrated clearly and recently in the cartel enforcement context. In 2003, the Antitrust Division revoked conditional leniency from Stolt-Nielsen S.A., a Norwegian company involved in a cartel within the bulk chemical shipping industry. The Division’s revocation was grounded in the allegations of co-conspirators who had been successfully prosecuted on the basis of evidence contained within Stolt-Nielsen’s leniency application to U.S. antitrust authorities. These co-conspirators alleged that Stolt-Nielsen had failed to take “prompt and effective action” to terminate its involvement in the conspiracy after it first discovered the anti-competitive activity. Following a series of civil proceedings challenging the revocation of leniency, the Division indicted Stolt-Nielsen for its self-reported Sherman Act violations. However, the district court dismissed the indictment, finding after an evidentiary hearing that the Division’s allegations that Stolt-Nielsen breached the leniency agreement were meritless,

105 Hammond, Summary Overview, supra note 49, at 5.
107 Id. at 614.
108 Id. at 614, 623.
109 Id. at 616.
110 Stolt-Nielsen had successfully sought to enjoin the Division from revoking leniency. Stolt-Nielsen S.A. v. United States, 352 F. Supp. 2d 553, 555 (E.D. Pa. 2005). However, the Third Circuit reversed the district court’s ruling, Stolt-Nielsen S.A. v. United States, 442 F.3d 177 (3d Cir. 2006), after which the Division filed its indictment.
111 Stolt-Nielsen, 524 F. Supp. 2d at 615.
and therefore that the Division’s revocation of leniency and subsequent prosecution was improper.\textsuperscript{112} Importantly, the court found that the testimony of the Division’s witnesses—the cooperating co-conspirators—lacked credibility.\textsuperscript{113}

First, the court noted that each of the witnesses had a strong motive to lie to the Division in exchange for leniency in their own criminal sentences.\textsuperscript{114} Second, when confronted on cross-examination, the witnesses’ were not credible and were repeatedly impeached.\textsuperscript{115} The Division witnesses’ testimonies were self-contradictory and riddled with material misstatements of fact, which appeared both in the live testimony and in sworn grand jury statements.\textsuperscript{116} Absent the right to cross-examine these government witnesses, it is unlikely that Stolt-Nielsen would have been able to prevail in its motion, and could have faced extraordinary fines and prison terms based almost entirely on the false testimony of its co-conspirators.\textsuperscript{117} The case of Stolt-Nielsen highlights the acute danger of using the out-of-court statements of co-conspirators—even when sworn—to establish cartel violations, as well as the unassailable value of cross-examination in the antitrust context.

\textsuperscript{112} Id. at 628.
\textsuperscript{113} Id. at 623.
\textsuperscript{114} Id. ("Each witness . . . had a strong motive to seek leniency from the Division and to retaliate against a competitor that had implicated him in a criminal conspiracy.").
\textsuperscript{115} Id. at 623-27. When discussing the credibility of co-conspirator Hugo Finlay, the court noted that he "was impeached repeatedly with prior inconsistent sworn testimony," first denying knowledge of the conspiracy “despite the fact that he had actively participated in it.” Id. at 624. In addition, when confronted about allegations he had made with regard to an anticompetitive quid pro quo agreement, Finlay “conceded on cross-examination that he had no personal knowledge of such a quid pro quo.” Id. With respect to another co-conspirator witness, Jarle Haugsdal, the court noted that Haugsdal had “provided repeated false accounts” about his company’s role in the conspiracy in his plea agreement, and that his “sworn grand jury declaration . . . was replete with material misstatements of fact.” Id. at 624-25. Furthermore, “[w]hen confronted with the inconsistencies, Haugsdal was uncertain how or by whom his declaration was prepared.” Id. at 625. Moreover, while cooperating witness Erik Nielsen testified “[o]n direct examination . . . that Stolt-Nielsen continued to participate in the conspiracy after [it discovered the anticompetitive conduct], when confronted with examples of vigorous post-March 2002 competition, Nielsen conceded that it was not ‘business as usual,’ [i.e. anticompetitive,] and repeatedly disavowed familiarity with the business,” stating that he “was not involved at all,” and “not familiar at all with these matters.” Id. at 626. Finally, and most incredibly, one cooperating witness, Bjorn Sjaastad, stated that "he was not aware that his conduct was illegal until he read [a newspaper] article reporting on antitrust violations in the parcel-tanker industry." Id.
\textsuperscript{116} See id. at 623-27.
\textsuperscript{117} Id. at 623.
II. The European Commission Approach to Cartel Enforcement

Similarly to the United States, the EU has had a significant history of cartel enforcement, animated from the start by the need to protect the European common market. Cartel enforcement falls within the purview of the EC, an “integrated” administrative agency within the EU that performs its own investigations, regulatory enforcement, and adjudications. Investigations typically begin either with a customer complaint sent to the EC, ex officio, or through an application for leniency. The investigatory stage of cartel enforcement in the EC is conducted by the Directorate General for Competition (“DG Competition”). These investigations use aggressive techniques inspired by American criminal cartel enforcement, and increasingly rely on oral evidence gathered in leniency applications and raids on targeted individuals and businesses. Upon a determination that there is sufficient evidence to prosecute undertaking party, the DG Competition case team prepares a “Statement of Objections,” which outlines the factual bases for the violation alleged.

The filing of a Statement of Objections triggers a series of procedural “rights of defense” for the alleged infringers, including the “right of access” to the DG Competition case file.

---

118 Directorate General for Competition, European Commission, http://ec.europa.eu/dgs/competition/mission/ (last visited Apr. 20, 2010) (“The mission of the Directorate General for Competition (DG Competition) is to enable the Commission to make markets deliver more benefits to consumers, businesses and the society as a whole, by protecting competition on the market and fostering a competition culture. We do this through the enforcement of competition rules and through actions aimed at ensuring that regulation takes competition duly into account among other public policy interests.”).


121 Id. at 6-7 (“The Commission may also open a case on its own initiative (ex officio), for instance when certain facts have been brought to its attention, or further to information gathered in the context of sector enquiries, informal meetings with industry or the monitoring of markets, or on the basis of information exchanged within the European Competition Network . . . .”).

122 Id. at 7.

123 BEST PRACTICES, supra note 120, at 6-17; Aslam & Ramsden, supra note 7, at 61; VAN BAEL & BELLIS, COMPETITION LAW OF THE EUROPEAN COMMUNITY § 10.10 (2004).

124 Forrester, supra note 7, at 833.

125 BEST PRACTICES, supra note 120, at 18.
and the “right to be heard” by the Commission. The primary catalysts for these procedural developments have come from common law Member States, the United States, and the European appellate courts, specifically the Court of First Instance (“CFI”) and the European Court of Justice (“ECJ”). Ultimately, these procedural rights were formalized in EC Regulations 1/2003 and 773/2004, which represent a convergence of the procedural safeguards in the EU and those of adversarial criminal justice systems in the United States and elsewhere. However, despite the fact that the EC has continually improved the defensive rights of accused entities, it has neglected to impose one fundamental and necessary right: the right of confrontation.

A. Competition Law and Procedure in the EU

The EU prohibits anticompetitive conduct in Article 101 of the Treaty on the Functioning of the European Union (“TFEU”). Specifically, Article 101(1) prohibits “all agreements between undertakings, decisions by associations of undertakings and concerted practices which may affect trade between Member States and which have as their object or effect the prevention, restriction or distortion of competition within the internal market.” An “undertaking,” for the purposes of Article 101, covers a broad array of entities,


127 With the passage of the Lisbon Treaty, the CFI is now entitled the “General Court,” however, because the pre-Lisbon decisions and authorities cited here refer to the “CFI,” this note will refer to the court by its former name. See Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community, signed at Lisbon, 13 December 2007, art. 2, O.J. C 306/1, at 43 (2007).

128 EC decisions are subject to judicial review by the CFI on issues of fact and law, and further subject to legal review before the ECJ. Carl Baudenbacher, Judicialization of European Competition Policy, in ANNUAL PROCEEDINGS OF THE FORDHAM CORPORATE LAW INSTITUTE: INTERNATIONAL ANTITRUST LAW & POLICY 354 (B. Hawk, ed. 2003); Asimow & Dunlop, supra note 126, at 146.

129 Asimow & Dunlop, supra note 126, at 144.


131 Treaty on the Functioning of the European Union, supra note 8, at 88.

132 Id.
including all legal and natural persons participating in economic or commercial activity.\textsuperscript{133}

Regulation 1/2003 sets forth the rules related to the enforcement of Article 101, and delegates enforcement responsibility both to the EC through the DG Competition and to Member States of the European Union.\textsuperscript{134} Under this regulation, the EC has the authority to “impose on [infringing undertakings] any behavioral or structural remedies which are proportionate to the infringement committed and necessary to bring the infringement to an end.”\textsuperscript{135} The EC is also authorized through Regulation 1/2003 to impose penalties—specifically and exclusively fines\textsuperscript{136}—on offending undertakings.\textsuperscript{137} These fines have exceeded 13 billion euros over the last 15 years.\textsuperscript{138}

Finally, under Article 230(1) of the EC Treaty, decisions made


\textsuperscript{134} Regulation No. 1/2003, art. 11-24, O.J. L 1/1, at 10-18 (2003) [hereinafter Regulation 1/2003].

\textsuperscript{135} Id. at 9. However, before the Commission may issue a decision regarding an undertaking, it must consult the Advisory Committee on Restrictive Practices and Dominant Positions (the “Advisory Committee”), which is comprised of representatives of Member States and prepares opinions on draft Commission decisions. The Commission must consider the opinions of the Advisory Committee with the “utmost account” in preparing a final decision. Id.

\textsuperscript{136} Baudenbacher, \textit{supra} note 128, at 354. In fact, the Commission has asserted its non-criminal character expressly “because it did not want Article 6 of the European Human Rights convention to be applied to it.” Id. Under Article 6 of the ECHR:

\begin{quote}
[e]veryone charged with a criminal offence has the following minimum rights: (a) to be informed promptly . . . of the nature and cause of the accusation against him; (b) to have adequate time and facilities for the preparation of his defence; (c) to defend himself in person or through legal assistance of his own choosing or, if he has not sufficient means to pay for legal assistance, to be given it free when the interests of justice so require; (d) to examine or have examined witnesses against him and to obtain the attendance and examination of witnesses on his behalf under the same conditions as witnesses against him; [and] (e) to have the free assistance of an interpreter if he cannot understand or speak the language of the court.
\end{quote}

Slater et al., \textit{supra} note 7, at 4 n.10.

\textsuperscript{137} The regulation states that the Commission may impose fines up to 10% of the undertaking’s “total turnover” in the previous business year in the event that the undertaking has infringed Article 101 or contravened an “interim measure” imposed upon the undertaking, or if the undertaking has failed to comply with a commitment it has made in response to a preliminary finding of infringement. Regulation 1/2003, \textit{supra} note 134, art. 23, at 16-17.

by the EC are subject to appellate judicial review by the CFI and ECJ.\textsuperscript{139}

B. Sources of Oral Evidence in EC Cartel Cases

The EC utilizes a multifarious approach to investigation and evidence-gathering that is heavily influenced by the American model.\textsuperscript{140} While in the past the Commission’s evidence was limited, to a large extent, to documentary evidence, its modern approach increasingly relies on oral evidence gathered from (1) the EC Leniency Program and (2) its investigations. As explained below, this modern approach bears heavily on the adequacy of EC procedural fairness.

1. The EC Leniency Program and Paperless Applications

The first source of oral evidence in EC cartel cases arises in the form of oral applications for leniency, which were introduced in the EC’s 2002 revision to its 1996 Leniency Program.\textsuperscript{141} The program was revised in 2002 and then again in 2006 “because [the EC] wanted it to be more attractive and . . . closer to the American program.”\textsuperscript{142} The current leniency program—outlined in the EC’s “Leniency Notice”—provides conditional immunity from fines to the first undertaking that approaches the Commission with evidence of cartel activity.\textsuperscript{143} This evidence must be in the form of a corporate statement that is sufficient to allow the Commission (1) to “carry out a targeted inspection in connection with the alleged cartel” (“Point 8(a) Immunity”),\textsuperscript{144} or (2) to “find an infringement of

\textsuperscript{139} See Holmes & Girardet, supra note 133, at 67.
\textsuperscript{140} International Cartels Roundtable, supra note 19, at 100.
\textsuperscript{141} Joshua, supra note 26, at 16.
\textsuperscript{144} Id. ¶ 8(a). In order to qualify for Point 8(a) Immunity, the applying undertaking must approach the Commission before it has “sufficient evidence to adopt a decision to carry out an inspection in connection with the alleged cartel or had already carried out such an inspection.” Id. ¶ 10. Furthermore, the leniency notice provides that an applicant for 8(a) Immunity must provide the Commission with (a)
Article [101 TFEU] in connection with the alleged cartel” (“Point 8(b) Immunity”). Additionally, the Leniency Notice provides fine reduction for subsequent applicants that provide “significant added value” to the EC’s investigation, defining “added value” with reference to “the extent to which the evidence provided strengthens, by its very nature and/or its level of detail, the Commission’s ability to prove the alleged cartel.” As in the United States, all applicants must meet a series of preconditions designed to ensure cooperation with the EC, evidence preservation, and effective and immediate withdrawal from the conspiracy.

Corporate statements made for leniency purposes may be submitted either in written or oral form. Oral statements were first allowed in the 2002 revision to the Leniency Notice, a modification encouraged by American lawyers as a means of protecting applicants from U.S. civil antitrust lawsuits brought by customers, indirect purchasers, consumers, and others. In evidence in the form of a “corporate statement” and (b) other evidence, including contemporaneous evidence of the conspiracy available to the applicant. Id. ¶ 9.

Id. ¶ 8(b). Where the EC has already initiated an investigation into cartel conduct without recourse to a Point 8(a) Immunity applicant, a cartel participant may qualify for Point 8(b) Immunity if it is “the first to provide contemporaneous, incriminating evidence of the alleged cartel as well as a corporate statement containing the kind of information . . . which would enable the Commission to find an infringement of Article [101 TFEU].” Id. ¶ 11.

Id. ¶ 25, at 20.

Id. ¶ 12, at 18. First, the applicant must cooperate “genuinely, fully, on a continuous basis and expeditiously from the time it submits its application throughout the Commission’s administrative procedure.” Id. ¶ 12(a). Second, the applicant must have ended “its involvement in the alleged cartel immediately following its application, except for what would, in the Commission’s view, be reasonably necessary to preserve the integrity of the inspections.” Id. ¶ 12(b), at 19. Third, the applicant “must not have destroyed, falsified or concealed evidence of the alleged cartel nor disclosed the fact or any of the content of its contemplated application, except to other competition authorities.” Id. ¶ 12(c). The last requirement outlined in the Leniency Notice concerns the applicant’s role within the cartel; namely, the applicant must not have coerced any other co-conspirator to join or remain a part of the cartel. Id. ¶ 13.

Id. ¶ 9(a), at 18 n.2.

Bertus van Barlingen & Marc Barennes, The European Commission’s 2002 Leniency Notice in Practice, 3 EUR. COMPETITION NEWSL. 6, 8 (2005) (“[T]he Commission allows [oral leniency applications] . . . in order to ensure that by making an application under the Commission’s Leniency Notice, undertakings are not worse off than non-cooperating cartel members in respect of civil procedures for damages.”); Joshua, supra note 26, at 14 (explaining that “[t]he exposure arises from the mathematical certainty that any announcement of an antitrust investigation in the US will trigger a welter of expensive and burdensome treble damages claims.”); 2003 FORDHAM CORP. L. INST 120 (ed. Barry Hawk 2004) (As Olivier Guersent, the former DG Competition noted, “[W]hat we want to protect from discovery is this very incriminating document that the applicants assemble for the Commission, and that is basically a roadmap to these documents. . . . We do believe these types of document [sic] should not be disclosed in civil trials because if they are, then they unbalance the
this so-called “paperless” procedure, the applicant or its attorneys may recite the relevant facts required for leniency onto a tape, which becomes “original evidence” added to the investigation file. After the applicant delivers its oral testimony, it will be “granted the opportunity to check the technical accuracy of the recording, which will be available at the Commission’s premises[,] and to correct the substance of their oral statements within a given time limit.” The EC shields the oral statement transcripts from discovery during the course of the investigation, and maintains unattested transcripts as an “internal resource” such that they are not discoverable either with the Statement of Objections or to civil plaintiffs. However, while protected from civil discovery, these oral statements are considered “evidence” in cartel cases against co-conspirators, a policy recognized both by the CFI and by the EC. Furthermore, and as more fully explicated below, Regulation No. 1/2003 allows EC competition case teams to “interview any natural or legal person who consents to be interviewed for the purpose of collecting information relating to the subject-matter of an investigation,” including interviews during the course of the leniency application process. However, these interviews are strictly voluntary, and the EC may not require individuals to give testimony under oath.

---

150 van Barlingen & Berennes, supra note 149, at 10. Originally, however, the oral statements were designed solely to gather “dawn raid sufficient” information. Julian Joshua, The European Cartel Enforcement Regime Post-Modernization: How is it Working?, 13 GEO. MASON L. REV. 1247, 1262 (2006) (“In the EC, however, the oral process has slipped from the provision of information sufficient to trigger a raid—that is, the information gathered is used as no more than a road map to the evidence—to being treated by the Commission as itself conclusive evidence of the violation.”).

151 Commission Notice, supra note 143, ¶ 32.

152 Joshua, supra note 26, at 15.

153 van Barlingen & Berennes, supra note 149, at 8; Joshua, supra note 26, at 17.

154 JFE Eng’g Corp. v. Comm’n, Joined Cases T-67/00, T-68/00, T-71/00 & T-78/00, 2004 E.C.R. II-2514, 2587 ¶ 192 (noting that “no provision or any general principle of Community law prohibits the Commission from relying, as against an undertaking, on statements made by other incriminated undertakings”). Specifically related to oral statements, the court—deciding the issue based on the 1996 Leniency Notice—stated that informational evidence “need not necessarily be provided in documentary form.” Tokai Carbon Co. Ltd. v. Comm’n, Joined Cases T-236/01, T-239/01, T-244/01 to T-246/01, T-251/01 & T-252/01, 2004 E.C.R. II-1181, ¶ 431.

155 van Barlingen & Berennes, supra note 149, at 9.


Thus, the leniency program involves two species of oral evidence, the applications and the interviews, neither of which is sworn or subject to cross-examination.\footnote{Lenaerts, supra note 157, at 1468-69 (“The Commission is entitled to rely on these submissions in order to establish the existence of an infringement. Since the principle is that the evaluation of evidence should be unfettered, this type of evidence is admissible under Community competition law. However, in the administrative procedure, the Commission does not have the power ‘to compel persons to give evidence under oath.’”) (quoting Rhône-Poulenc SA, 1991 E.C.R. II-867, II-954) (Vesterdorf, J.).}

2. Investigation and Oral Interviews

The second source of oral evidence in EC cartel cases arises from investigations into alleged violations of European competition laws, as authorized under Regulation 1/2003.\footnote{Regulation 1/2003, supra note 134, art. 17-22, at 13-16.} First, where the EC has reason to believe that there is a distortion or restriction in the common market, it may investigate any specific industry or market with respect to the apparent distortion or restriction, and may request any necessary information from undertakings, including information related to “all agreements, decisions and concerted practices.”\footnote{Id. art. 17, at 13.} Additionally, the EC may, by request or decision, “require undertakings and associations of undertakings to provide all necessary information” to assist with the investigation.\footnote{Id. art. 18(1), at 13.} In the event that an undertaking provides false or misleading information, or does not provide the requested information within the prescribed time limit, the undertaking may face a fine of up to 1% of its “total turnover in the preceding business year . . . .”\footnote{Id. art. 23(1)(a)-(b), at 16.}

In practice, a team led by a DG Competition case manager will conduct the investigation into the alleged violations, which may include either a document request or a “dawn raid” on a targeted business.\footnote{A dawn raid is an early morning, “on-the-spot” investigation of a business office or private residence conducted by EC officials—at times with the assistance of law enforcement from the member state—pursuant to Regulation 1/2003 Articles 18 through 21. See Asimow & Dunlop, supra note 126, at 156; Imran Aslam & Michael Ramsden, EC Dawn Raids: A Human Rights Violation?, 5 COMPETITION L. REV. 61, 65-67 (2008); Regulation 1/2003, supra note 134, art. 18-21, at 13-16.} Increasingly, investigations are spurred by leniency applications, which require “dawn raid sufficient” evidence.\footnote{Joshua, supra note 9, at 4.} During the course of
an investigation, “the Commission may interview any natural or legal person who consents to be interviewed for the purpose of collecting information relating to the subject-matter of an investigation.” Further, the case team may ask individuals to provide voluntary oral explanations of the documentary evidence it collects in connection with the dawn raids, which may include searches of businesses, residences, and automobiles. The testimonial evidence gathered in these interviews and explanations may be used as evidence in the Statement of Objections, and is not sworn or subject to cross-examination.

C. Procedural Safeguards in the EC

In the event that the DG Competition case team discovers violations in the course of an investigation, it may seek approval from the Legal Service and the Competition Commissioner, as well as a committee of European Member States, to outline a “Statement of Objections,” which includes (1) the charges against the undertaking and (2) the time limit for the undertaking to respond to the allegations. Upon the filing of a Statement of Objections, the undertaking is afforded a series of procedural rights, including (1) the right to access the Commission’s file against the undertaking and (2) the right to be heard in writing or at an oral hearing.

1. The Right of Access to File

Parties subject to a Statement of Objections are “entitled to have access to the Commission’s file,” excepting business secrets and confidential information such as internal documents prepared by the EC and Member States of the EU. The right of access was first articulated in Solvay v. Commission, in which the CFI explained the purpose of the right; namely, “to enable addressees of statements of objections

166 Lenaerts, supra note 157, at 1469.
168 Id.
169 Id.
170 BEST PRACTICES, supra note 120, at 18-19; Asimow & Dunlop, supra note 126, at 156.
171 Id. at 36-39.
172 Regulation 1/2003, supra note 134, art. 27(2), at 19.
to examine evidence in the Commission’s file so that they are in a position effectively to express their views on the conclusions reached by the Commission in its statement of objections on the basis of that evidence.” Furthermore, the Court outlined the “equality of arms” principle, which states that the information available to the Commission and the undertaking at issue must be equal.

2. The Right to Be Heard

The right to be heard was first outlined in *Transocean Marine Paint Ass’n v. Commission*, where the CFI noted that “a person whose interests are perceptibly affected by a decision taken by a public authority must be given the opportunity to make his point of view known.” Later, in *Hoffmann-La Roche & Co. v. Commission*, the CFI generalized the procedural rights of undertakings, stating that “the undertakings concerned must have been afforded the opportunity during the administrative procedure to make known their views on the truth and relevance of the facts and circumstances alleged and on the documents used by the Commission to support its claim” of infringement. The right to be heard is now formalized in Regulation 1/2003, which requires that, prior to making a decision with respect to an alleged violation of TFEU Article 101, the Commission must give the alleged violators an “opportunity to be heard” on the matters set forth in the Statement of Objections.

Oral hearings are conducted by a Hearing Officer independent of the investigative team. The CFI has held that

---

174 Shipley, supra note 1, at 40.
177 The hearing procedures themselves are more fully explained in a subsequent, clarifying regulation, Regulation 773/2004, which states that the parties must be given an “opportunity to develop their arguments at an oral hearing” prior to the Commission consulting with the Advisory Committee. Commission Regulation No. 773/2004, art. 12, 2004 O.J. (L 123) 21 (EC) [hereinafter Regulation 773/2004]; Regulation 1/2003, supra note 134, art. 12, at 19; Joined Cases 100-103/80, SA Musique Diffusion Francaise v. Comm’n, 1983 E.C.R. 1825, 1881-84; Shipley, supra note 1, at 36.
178 Regulation 773/2004, supra note 177, art. 14(1), at 21 (“Hearings shall be conducted by a Hearing Officer in full independence.”); Committee Decision of 23 May 2001 on the Terms of Reference of Hearing Officers in Certain Competition Proceedings, 2001 O.J. (L 162) 21 (“The conduct of administrative proceedings should therefore be
these hearings are “adversarial” in nature, and each hearing generally follows these steps: (1) the Commission presents its case, (2) the defending parties and any relevant third parties present their case in response to the statement of objections, (3) representatives of the Member States and the Commission may ask questions regarding the arguments presented in the defending parties’ presentations, and (4) the Hearing Officer may provide the parties with an opportunity to make brief, concluding remarks relating to issues previously discussed during the hearing.\(^{79}\) Additionally, EC competition hearings are recorded.\(^{180}\)

During its presentation to the Commission, a defending party has the right to submit testimony in response to the Statement of Objections, may call on its own fact and expert witnesses, and may submit questions to the Hearing Officer, who then has the discretion to put the questions to Commission witnesses.\(^{181}\) However, unlike in adversarial systems such as the United States, parties in hearings before the European Commission “do not have the right to cross-examine the Commission, other parties (co-defendants) or third persons whose testimony is heard at the hearing,” and “[t]he Commission does not cross-examine the parties, but it may question them after the presentation with respect to their oral

---

entrusted to an independent person experienced in competition matters who has the integrity necessary to contribute to the objectivity, transparency and efficiency of those proceedings.”). As the EC states:

The hearing officer brings a new pair of eyes to trade proceedings and is fully impartial. As an official of DG Trade experienced in trade issues, he has a thorough knowledge of the system from the inside, but is not involved in ongoing investigations or conducting trade proceedings himself. The hearing officer is independent from the Commission investigators and receives no instructions from them about his substantial role.


\(^{181}\) Forrester & Komninos, *supra* note 130, at 61.
submissions.” After the hearing, the Hearing Officer is responsible for preparing an interim and final report for the College of Commissioners, which then decides whether to impose penalties on the defending undertaking. Thus, while “[t]he procedure in [EC competition cases] comes very close to the full evidentiary hearing of the type required by . . . Goldberg v. Kelly,” the lack of cross-examination—especially given the increasing use of oral evidence—highlights the inadequacy of the present safeguards.

III. TRENDS IN EC CARTEL ENFORCEMENT TOWARD A CRIMINAL AND ADVERSARIAL MODEL, AND THE FAILURE TO MAINTAIN ADEQUATE PROCEDURAL SAFEGUARDS

The current state of EC cartel enforcement, as outlined above, represents the latest stage in an “important evolutionary process” with respect to its administrative procedure—one that has brought it strikingly close to the American model and now demands an attendant right to confrontation. First, as investigations have become more aggressive and fines have increased, the EC cartel enforcement regime has become increasingly “criminalized.” Second, EC procedures have become increasingly “judicialized,” adopting structures to provide independence of decision-making and to sever the oft-criticized role of the EC as investigator, prosecutor, and judge. Third, as the EC has developed its procedures, there has been an increase in procedural rights.

---

182 Id. at 63.
183 These reports relate to both procedural and substantive issues, including whether the rights of defense were respected in the hearing and the Hearing Officer’s substantive observations on the course of the proceedings. DG COMPETITION, THE HEARING OFFICERS, supra note 179, at 16-17.
186 Asimow & Dunlop, supra note 126, at 143-44; Philip Marsden & Peter Whelan, Re-Examining Trans-Atlantic Similarities and Divergences in Substantive and Procedural Competition Law, 10 SEDONA CONF. J. 23, 23 (2009).
187 Forrester, supra note 7, at 834 (“We have come a long way since the early days of EC competition law, when cases were carefully picked, decisions were few and penalties were modest. The regime had been set up in 1962, it was a relatively obscure topic, and the fines were not confiscatory though they could, in theory, be painful.”); see also Aslam & Ramsden, supra note 7, at 61-62; Slater, Thomas & Waelbroeck, supra note 7, at 4, 11.
188 Forrester, supra note 7, at 830; see also Asimow & Dunlop, supra note 126, at 157; Lenaerts, supra note 157, at 1474; Shipley, supra note 1, at 14.
afforded to private parties before the EC. However, the advent of oral leniency applications in the EC in 2002 and the increasing use of oral evidence in EC competition proceedings raise a number of questions about the sufficiency of extant procedural protections. The credibility of oral leniency applicants is especially lacking because those individuals and corporations have a strong interest in embellishing the role of co-conspirators while minimizing the extent of their own participation in the cartel. Furthermore, the lack of procedures for confrontation and cross-examination in EC proceedings threatens the integrity of the entire fact-finding process, since “hearsay accounts given by lawyers fall short of any generally accepted evidential standards, especially if they are the only proof adduced by the Commission.” While some commentators argue that the protections in the European Union adequately provide procedural fairness, others are not convinced. To secure procedural rights for corporations, and in light of the implementation of the oral leniency application, this note proposes that the EC should adopt a Crawford v. Washington-based approach to confrontational rights. Adapting this approach will serve a vital, legitimating function in a system where policy changes have outpaced procedural adaptations.

A. Criminalization of EC Cartel Enforcement

The EC’s cartel enforcement regime has appropriated elements of American cartel enforcement since its inception.

---

189 Francesca Bignami, Creating European Rights: National Values and Supranational Interests, 11 COLUM. J. EUR. L. 241, 279 (2005); Lenaerts, supra note 157, at 1474; Wils, supra note 19, at 128.
190 Joshua, supra note 9, at 1; Lenaerts, supra note 157, at 1474-94 (describing existent procedural protections in European Community Competition Law).
191 Forrester, supra note 7, at 833 (“Leniency applicants have a clear interest in showing serious wrongdoing by their competitors. The juicier the information they provide, the more chance they have of being deemed to have provided ‘significant added value’ and obtaining a reduction in their fine, or total immunity; and, in the process, they will have created big trouble for their competitors.”); Joshua, supra note 9, at 7.
192 Joshua, supra note 9, at 6.
193 Id. at 16.
194 Shipley, supra note 1, at 48.
195 Joshua, supra note 9, at 1.
196 Crawford, 541 U.S. at 68.
197 Nicholas Green QC, The Road to Conviction—The Criminalisation of Cartel Law, in 2003 FORDHAM CORP. L. INST. 000, 28 (B. Hawk, ed. 2004) (though
When the European Union crafted the Treaty of Rome\textsuperscript{198} in 1957, it “took as its inspiration for Articles 85 and 86EC, sections 1 and 2 of the Sherman Act.”\textsuperscript{199} This influence extends to the means of investigation employed by the EC in cartel enforcement, as well as the ever-increasing fines that it levies on violators.\textsuperscript{200}

1. Criminalized Investigations

Over time, the European Commission has adopted many of the investigatory techniques utilized by criminal enforcement regimes like the United States, and sometimes utilizes powers that go beyond those criminal systems.\textsuperscript{201} First, the most invasive investigatory power of the European Commission is the “dawn raid.”\textsuperscript{202} Under Regulation 1/2003, the Commission has sweeping authority to conduct searches of businesses, private residences, and private automobiles, subject only to minimal oversight by national courts.\textsuperscript{203} Furthermore, when the Commission effectuates these searches, it often utilizes the law enforcement mechanisms—including search warrants—and police personnel of the EU Member States.\textsuperscript{204} Some commentators have intimated that these techniques were inspired by the U.S. antitrust authorities’ successful use of criminal search warrants to retrieve business records from the homes of cartel participants.\textsuperscript{205} Irrespective of their source, however, it is clear that the increasingly aggressive EC investigations are closely aligned with criminal law enforcement techniques used in the United States.\textsuperscript{206}

Second, the Commission adopted a leniency program—a tool that had traditionally been reserved for only the most

---


\textsuperscript{199} Green QC, supra note 197, at 28.

\textsuperscript{200} Chavez, supra note 30, at 824-25; Wils, supra note 19, at 14.

\textsuperscript{201} See Aslam & Ramsden, supra note 163, at 65-67.

\textsuperscript{202} Id.

\textsuperscript{203} Id.; see also Holmes & Girardet, supra note 133, at 68.


\textsuperscript{205} Chavez, supra note 30, at 825.

\textsuperscript{206} Aslam & Ramsden, supra note 7, at 64-67.
serious criminal conspiracies—in large measure as a result of the successful American experiment with its program in the criminal context.\textsuperscript{207} The EC further undertook to bring its leniency program in line with the US model through its 2002 and 2006 revisions, in which the most notable amendments involved the incorporation of confidentiality provisions and the oral leniency procedure.\textsuperscript{208} These provisions, designed to protect against disclosure in American civil litigation, brought the program “closer to the U.S. regime that is based on oral statements rather than written documents.”\textsuperscript{209} However, while the American system safeguards defendants from the use of these oral statements in trial through its entrenched procedural defense rights, the oral statements in the EC procedure only seemed to create “a number of new problems/issues,” including: (1) “issues related to the evidentiary value of recorded oral statements in Community antitrust procedures” and (2) “issues related to the modalities of exercise of the rights of defence of the other cartel members.”\textsuperscript{210}

2. Penalties Exhibit Criminal Qualities

The recent modernization efforts in EC cartel enforcement were “triggered by the need to restore the effectiveness of the fight against secret unlawful agreements.”\textsuperscript{211} One of the most significant ways the EC did this was to adopt “a considerable increase in the level of fines imposed by the Commission.”\textsuperscript{212} A CFI judge has noted that EC fines “have a criminal law character,” such that, increasingly “parties’ submissions can only be understood with the help of the terminology and concepts used in criminal law and procedure.”\textsuperscript{213} While EC decisions against cartel participants between 1969 and 1995 totaled 3.329 million euros (comprising eighty cartel decisions), enforcement between 1996 and the end

\begin{footnotes}
207 \textsuperscript{} Wils, \textit{supra} note 19, at 127.
208 \textit{International Cartels Roundtable, supra} note 19, at 100.
209 Olivier Guersent, \textit{The Fight Against Secret Horizontal Agreement in the EC Competition Policy}, \textit{in} 2003 FORDHAM CORP. L. INST. 43, 54 (B. Hawk, ed. 2004); \textit{International Cartels Roundtable, supra} note 19, at 100.
210 Guersent, \textit{supra} note 209, at 54.
211 \textit{Id.} at 45.
212 \textit{Id.} at 47; \textit{see also} Marsden & Whelan, \textit{supra} note 186, at 23.
\end{footnotes}
of 2003 totaled 3.45 billion euros (comprising thirty decisions), with fines exceeding 9 billion euros for the 2004 to 2009 timeframe (comprising 243 decisions).214 This trend toward astronomical fines exhibits the increasingly punitive and quasi-criminal character of EC cartel enforcement, heightening the “relevance of general principles of substantive criminal law” and its attendant procedural safeguards.215

Furthermore, Regulation 1/2003 provides for indirect criminal penalties through the operation of the cartel enforcement regimes of the Member States. Under Article 5 of Regulation 1/2003, “[t]he competition authorities of the Member States shall have the power to apply Articles [101] and [102] of the [TFEU] in individual cases,” and they may take decisions “imposing fines, periodic penalty payments or any other penalty provided for in their national law.”216 As a result of this broad penalty authorization, Member States may levy criminal sentences on corporations and individuals alike, and such penalties have been extended in Ireland, Estonia, and the United Kingdom.217 Thus, while these penalties are not directly imposed or adjudicated by the Commission, their availability through the Member States suggest that the Commission is willing to accept the appropriateness of criminal consequences for violations of EC competition laws.

3. EC Cartel Hearings Exhibit Important Adversarial Qualities

At the procedural level, the trend in EU cartel enforcement has been one toward a quasi-criminal, adversarial system, replete with many—but not all—of the rights found in adversarial systems.218 First, although the CFI has declined to conclude that the Commission constitutes a “tribunal,” it contains many of the same characteristics.219 Namely, it (1) is established by law, (2) is permanent, (3) exercises compulsory

214 Guersent, supra note 209, at 48; see also EC Fine Statistics, supra note 138.
215 Lenaerts, supra note 157, at 1485.
216 Regulation 1/2003, supra note 134, art. 5, at 8-9; see also Wils, supra note 19, at 129.
217 Wils, supra note 19, at 130.
219 Forrester, supra note 7 at 831; Shipley, supra note 1, at 36.
jurisdiction, (4) carries out inter partes proceedings,\footnote{See sources cited supra note 219.} (5) applies the rule of law with evidentiary standards,\footnote{Guidance on Procedures, supra note 178, at 4. (“The Commission has to conduct its competition proceedings fairly and objectively while respecting the parties’ procedural rights. The Hearing Officers are, first of all, guardians of fair proceedings before the Commission. They safeguard the rights of defence of undertakings subject to proceedings relating to Articles 101 and 102 (ex-articles 81 and 82) as well as the procedural rights of . . . all . . . parties to the proceedings.”); Regulation No. 1/2003, supra note 134.} and (6) is, in certain important ways, independent.\footnote{DG COMPETITION, THE HEARING OFFICERS, supra note 179, at 5-6.; Baudenbacher, supra note 128, at 355; Regulation No. 773/2004, supra note 179, art. 14(1), at 21 (“Hearings shall be conducted by a Hearing Officer in full independence.”).} For example, the role of the Hearing Officer, who conducts EC competition hearings and files reports regarding compliance with procedural rules, severs the relationship between the investigating case team and the decision-making College of Commissioners.\footnote{DG COMPETITION, THE HEARING OFFICERS, supra note 178, at 6. The College of Commissioners is comprised of representatives of member states who are “completely independent in the performance of their duties[,] . . . neither seek[ing] nor taking[ing] instructions from any government or from any other body.” Id. Governance Statement of the European Commission 2, available at http://ec.europa.eu/atwork/synthesis/doc/governance_statement_en.pdf; see also Regulation No. 773/2004, supra note 177, art. 14(1), at 21.} This division of roles creates delineated prosecutorial and judicial functions, heightening the adversarial nature of the hearings and providing “an independent guarantor of the fundamental procedural rights of all parties.”\footnote{Baudenbacher, supra note 128, at 356.} Furthermore, the use of oral evidence in EC hearings has increased, which, while uncommon in inquisitorial hearings, stands as a hallmark of Western adversarial systems.\footnote{Id. at 355.} Based on the trend toward an adversarial system, it is not surprising that there has been a concomitant shift toward greater procedural rights in the EC.

B. Trend toward Greater Procedural Rights in EC Cartel Cases

Undertakings in EC competition hearings were afforded very little by way of procedural rights at the inception of the EC.\footnote{Asimow & Dunlop, supra note 126, 143.} However, over time there has been a significant increase in the “rights of defence” for targeted undertakings, including the right of “access to file” and the “right to be heard.”\footnote{At first, the accession of rights was inspired in part by intense pressure from the United Kingdom, which, having joined the EU, wanted to protect the fundamental rights of its citizenry. Id. at 143-46.}
However, these procedural rights have almost exclusively focused on rights exercised before and after the EC hearing, and do not include a right to cross-examine witnesses.\textsuperscript{228} Certainly, in a purely inquisitorial system, confrontation is much less necessary, both because an inquisitorial approach takes power out of the hands of the individual parties and because live, testimonial evidence bears much less weight than documentary evidence.\textsuperscript{229} In the EC, however, the increasing use of oral testimony and adversarial postures creates a disjuncture between its substantive policies and procedural protections.\textsuperscript{230} As such, where the starting place and crux of high-stakes cartel enforcement is increasingly unsworn, \textit{ex parte} oral testimony used explicitly as evidence of collusive conduct, procedural rights must include a right of confrontation.\textsuperscript{231}

\section*{IV. TOWARD A RIGHT OF CONFRONTATION FOR EC COMPETITION HEARINGS}

While trends in EC cartel enforcement have kept pace with global cartel enforcement with respect to investigation and punishment, the EC has not made parallel strides in the area of procedural protections.\textsuperscript{232} One fundamental right above all is lacking in EC competition procedure: the right of confrontation. The American constitutional right of a criminal defendant to confront and cross-examine witnesses against him provides a useful model for managing oral testimony that would bring necessary procedural fairness to the EC cartel enforcement regime.\textsuperscript{233}

\begin{itemize}
  \item \textsuperscript{228} Forrester, supra note 7, at 831-32; Joshua, supra note 9, at 6-7; Philip Lowe, Director General, DG Competition, Speech at the CRA Conference on Economic Developments in Competition Law: Due Process in Antitrust, at 7 (Dec. 9, 2009), available at http://ec.europa.eu/competition/speeches/text/sp2009_19_en.pdf.
  \item \textsuperscript{229} Lenaerts, supra note 157, at 1469.
  \item \textsuperscript{230} Forrester, supra note 7, at 831, 833.
  \item \textsuperscript{231} Joshua, supra note 26, at 16 (“Hearsay accounts given by lawyers fall short of any generally accepted evidential standards, especially if they are the only proof adduced by the Commission.”).
  \item \textsuperscript{232} Marsden & Whelan, supra note 186, at 40.
  \item \textsuperscript{233} For the purposes of this note, the Sixth Amendment Right of Confrontation, and not the Due Process Clause (or Federal Trade Commission administrative procedure), provides the correct analogy, based on the \textit{de facto} criminalization of cartel conduct in the EC and the increasingly formal and adversarial nature of EC competition hearings.
\end{itemize}
A. The Necessity of a Right of Confrontation for EC Competition Hearings

At the dawn of the adoption of the 2002 Leniency Notice, which first allowed for oral leniency applications, Emil Paulis, the former Director responsible for Policy and Strategic Support at DG Competition, stated that because “EU administrative proceedings [are] centered on an exchange of written arguments[,] . . . it does not serve the parties to cross-examine the Commission.” Further, in 1991 the ECJ held that—in the context of anonymous documents—the Commission need only perform “an overall assessment of a document’s probative value” to determine admissibility. However, because the evidence used in EC Competition hearings is now so heavily based on oral testimony, the document-based procedure of the past can no longer be credibly relied upon as a justification for the denial of confrontational rights.

First, the use of oral leniency applications in the EC as evidence of cartel conduct raises many of the hearsay and credibility issues against which cross-examination is designed to protect. Because cartels are inherently secretive, the possibility of ferreting out and establishing infringement absent a confession or oral explanation of incriminating documents is often remote. At the same time, however, there is a strong incentive for individuals and corporations to embellish the conduct of co-conspirators in the leniency application process in order to achieve “significant added value,” and there is no direct right to test the accuracy of those statements in EC oral hearings.

---

236 Joshua, supra note 150, at 1262.
237 Sklansky, supra note 72, at 1646 (“Introducing evidence of an out-of-court accusation from someone who never testifies raises some of the same concerns as examining a witness outside the defendant’s presence: in either case the defendant has no opportunity to cross-examine the accuser in front of the jury.”).
238 Van Barlingen & Barennes, supra note 149, at 9; Joshua, supra note 9, at 1, 3, 5-6.
239 Forrester, supra note 7, 833; Joshua, supra note 9, at 5-6. The incentive to embellish is two-fold. First, an undertaking applying for leniency has an incentive to provide evidence of its co-conspirator’s bad acts while downplaying its own role to avoid being viewed as a “ring-leader,” which would prevent it from receiving leniency.
In this way, the use of oral testimony and oral leniency applications as evidence absent a right to confrontation or cross-examination is analogous to the obtainment and use of the untested, *ex parte* statements of Lord Cobham in Sir Walter Raleigh’s trial—a practice viewed as anathematic to a fair trial in common law systems.\(^{241}\) Continuing the analogy to the Raleigh trial, in EC procedure, the leniency-seeking co-conspirator will have prepared an out-of-court confession under intense pressure and with dubious accuracy, and the accused has no procedural right to compel the presence of the co-conspirator or to cross-examine the accusatory witnesses.\(^{242}\) Additionally, in both cases the evil arises based in part upon an incompatible hybridization of adversarial and inquisitorial legal systems.\(^{243}\) Moreover, as demonstrated in the *Stolt-Nielsen* case, co-conspirators in the context of cartel enforcement may—and often will—lie in exchange for leniency or, in the case of also-rans, in retaliation for their co-conspirator’s admissions to competition authorities.\(^{244}\) In fact, the role of the government witnesses in the *Stolt-Nielsen* case is not unlike the position of the “also ran” leniency applicants in the EC, since both had or have a strong incentive to shade co-conspirator conduct in the least favorable light.\(^{245}\) Therefore, given the need to utilize oral evidence in cartel cases and the trend toward adversarial and criminal cartel enforcement in the EC, the fairest solution is to implement procedural reforms that would provide a right of confrontation in EC competition hearings.

Second, in the context of “also-ran” undertakings, there is a heightened incentive to provide as much evidence of anticompetitive behavior as possible, since the level of fine reduction is tied to both the quantity and quality of the information provided. 2006 O.J. (C 298) 17, 20 ¶ 23-26 available at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2006:298:0017:0022:EN:PDF.

\(^{240}\) Forrester, *supra* note 7, at 833; Joshua, *supra* note 9, at 6-7.

\(^{241}\) Sklansky, *supra* note 72.

\(^{242}\) Forrester, *supra* note 7, at 833; Sklansky, *supra* note 72.

\(^{243}\) In Raleigh’s trial, the adoption of inquisitorial practices in the English common law system prevented Raleigh from receiving a fair trial, whereas the current failure arises from the adoption of increasingly criminal and adversarial procedures without affording the accused fundamental rights. Crawford v. Washington, 541 U.S. 36, 44-45 (2004); Asimow & Dunlop, *supra* note 126, at 143; Forrester, *supra* note 7, at 833-36.


\(^{245}\) *Id.*
B. Proposals for Achieving Procedural Fairness in the EC

While the need for a right to confrontation within the EC is recognized by many, one remaining question is how the right should be introduced into the procedural scheme. First, some commentators call for an explicit criminalization of cartel conduct in the EC,\(^{246}\) which they view as a more effective means of both deterring cartel conduct and garnering procedural rights for targeted undertakings.\(^{247}\) Second, some contend that the inherent criminal nature of cartel conduct justifies an extension of the European Convention on Human Rights (“ECHR”) procedural protections, which include cross-examination, to the EC through the reviewing courts.\(^{248}\) Third, others suggest reform from within the EC as a part of a fundamental overhaul of EC procedure without direct reference to the ECHR.\(^{249}\) However, based on the EC apprehension of and distaste for these solutions, none are likely to be implemented.\(^{250}\) Instead, the best solution is to provide for a right of an undertaking subject to a Statement of Objections to confront and cross-examine any witness who provides ex parte oral testimony that the EC intends to use as evidence of a violation of the competition laws.


\(^{248}\) Slater et al., supra note 7, at 4.


\(^{250}\) Marsden & Whelan, supra note 186, at 25.
1. The European Convention on Human Rights

Some commentators have argued that EC cartel enforcement already includes significant hallmarks of criminal enforcement such that it should be subject to the procedural safeguards of the ECHR, which includes a right to cross-examination.\textsuperscript{251} According to this theory, because the ECJ “has always indicated its willingness to follow the case-law of the European Court of Human Rights (‘ECtHR’),”\textsuperscript{252} and because the ratification of the Lisbon Treaty requires EU compliance with the ECHR,\textsuperscript{253} the Commission’s procedures should square with Article 6 of the ECHR, leaving “the adjudicating function in antitrust cases to . . . a [competition specific] ‘judicial panel’ attached to the CFI.”\textsuperscript{254} Other commentators have argued for a less institutionally disruptive ECHR-based solution, under which the hearings will be mandatory in order to ensure ECHR compliance, and the hearing officer will be elevated to the role of finder of fact and law.\textsuperscript{255}

However, the call for judicial extension of the ECHR to EC cartel conduct is an ineffective and ultimately implausible solution. First, the European Union does not currently have the desire to formally criminalize cartel conduct.\textsuperscript{256} Further, because the hearings themselves were specifically designed to avoid ECHR application, any such modification will likely require outside introduction from the courts—as well as wholesale restructuring of EC competition procedure—which is impracticable.\textsuperscript{257} Additionally, the relevant courts have held that because the EC Commission is not a formal “tribunal” as defined under the ECHR, it is not subject to its procedural requirements.\textsuperscript{258} Finally, even if the levels of fines in EC hearings are “criminal” under the ECHR, commentators have noted that EC procedure is nonetheless compatible with Article 6 because the fines lie outside of the “hard core of criminal

\textsuperscript{251} Forrester, supra note 7, at 828-29; see also Slater et al., supra note 7, at 4.
\textsuperscript{252} Slater et al., supra note 7, at 3, 26 (noting that case law from the ECtHR has held that “[a]n oral, and public, hearing constitutes a fundamental principle enshrined in Article 6 § 1”)
\textsuperscript{253} Forrester, supra note 7, at 822.
\textsuperscript{254} Slater et al., supra note 7, at 46.
\textsuperscript{255} Forrester, supra note 7, at 841-43.
\textsuperscript{256} Marsden & Whelan, supra note 186, at 25.
\textsuperscript{257} Forrester, supra note 7, at 831.
\textsuperscript{258} Joined Cases 100-103/80, SA Musique Diffusion Francaise v. Comm’n, 1983 E.C.R. 1825, 1881-84.
law,” and therefore need not meet its stringent procedural requirements. Thus, a judicial extension of the ECHR is an inapt method for achieving procedural fairness in EC competition hearings.

2. Holistic Reform

Julian Joshua, former Deputy Head of the Cartel Unit of DG Competition, has seen first-hand the consequences of the 2002 EC procedural modernization efforts, and has developed a solution involving fundamental procedural reform within EC competition proceedings. In 1995, prior to the implementation of Regulation 1/2003 and the leniency program, he contended that the EC provided sufficient procedural safeguards. However, after the implementation of the leniency reforms in 2002, he began to advocate for a “bold and ‘holistic’ solution . . . encompassing the whole scope of the enforcement process,” both to improve the leniency program and to protect the “due process” rights of implicated parties.

First, where reliability is at issue, Joshua asserts that “contemporaneous documentary evidence” should be given more weight than statements that parties make during the course of proceedings, including oral corporate statements. Second, those statements made for the purpose of obtaining leniency should be viewed with particular caution, requiring a

259 Wouter P.J. Wils, The Increased Level of EU Antitrust Fines, Judicial Review and the ECHR, 33 WORLD COMPETITION L. & ECON. REV. 5, 17 (2010) (“Whereas the European Commission’s antitrust fining powers are not ‘criminal’ within the meaning of EU law, they are ‘criminal’ within the wider autonomous meaning of Article 6 ECHR. Inside the wider autonomous ECHR category of ‘criminal’, the requirements of Article 6 ECHR are different for, on the one hand, the ‘hard core of criminal law’; and, on the other hand, outside the hard core of criminal law. The European Commission’s antitrust fining powers . . . are outside the ‘hard core of criminal law’. Outside the hard core of criminal law, Article 6 ECHR allows for criminal penalties to be imposed, in the first instance, by an administrative or non-judicial body, that combines investigative and decision-making powers, provided that there is a possibility of appeal ‘before a judicial body that has full jurisdiction, including the power to quash in all respects, on questions of fact and of law, the challenged decision’. Article 6 ECHR thus allows the imposition of antitrust fines by the European commission and does not require any separation inside the Commission between investigative and decision-making functions, provided that the EU Courts, before which the addressees of Commission fining decisions can appeal, have ‘full jurisdiction, including the power to quash in all respects, on questions of fact and of law, the challenged decision’. . .”).

260 Joshua, supra note 218.

261 Joshua, supra note 26, at 21.

262 Joshua, supra note 9, at 7.

263 Id.
“high degree of corroboration by independent documentary evidence.”

Third, Joshua advocates for a procedure whereby oral evidence used to prove violations should be “reduced to writing in a formal witness statement, . . . signed by the individual, and endorsed and acknowledged by the company on whose behalf it is proffered.”

Fourth, he calls for an almost complete ban on anonymity for statements used to prove infringement.

Fifth, he simply states that “[d]eclarants should be available at the oral hearing for cross-examination by the parties incriminated.”

Sixth, he advocates for amendments to the EU regulations to penalize corporations and individuals for giving misleading testimony.

Lastly, he contends that oral leniency applications should serve merely as a “roadmap” for DG Competition investigations, and not as evidence of infringement.

Such a comprehensive overhaul in EC competition procedure, however, is unworkable and unlikely. First of all, as with the ECHR reform approach mentioned above, the EC currently lacks an appetite for such sweeping reform. Instead, a more palatable solution would provide real advancements in an incremental manner.

Second, the holistic reforms that Joshua proposes with respect to admitting oral hearsay testimony, based on a ‘reliability’ standard, contain the same dangers of subjectivity in the Ohio v. Roberts approach that the Supreme Court so persuasively rejected in Crawford v. Washington. In fact, the zealfulness with which competition

---

264 Id.
265 Id.
266 Id.
267 Id.
268 Id.
269 Joshua, supra note 9, at 7.
270 Id.
271 Neelie Kroes, European Comm’r for Competition Policy, Antitrust and State Aid Control-The Lessons Learned 5, Address at the 36th Annual Conference on International Antitrust Law and Policy (Sept. 24, 2009), available at http://ec.europa.eu/competition/speeches/index_theme_6.html. (“I agree again with Christine Varney here—as enforcers we do have ‘special responsibility’ to ensure a fair and transparent process. But the great weight of evidence says we meet this responsibility.”).
272 Marsden & Whelan, supra note 186, at 35.
273 Forrester, supra note 7, at 840.
274 541 U.S. 36. Although Joshua’s solution differs from the Ohio v. Roberts approach by virtue of its reliance on contemporaneous documentary evidence supporting reliability, the tendency in cartel enforcement in cartel cases to view each price quote and email through a conspiratorial lens means that co-conspirators can easily paint innocuous transmissions as severe violations. See, e.g., United States v. Stolt-Nielsen S.A., 524 F. Supp. 2d 609, 619-20 (E.D. Pa. 2007).
authorities have relied upon apparently reliable—but ultimately false—oral evidence, as evident in the *Stolt-Nielsen* case, reveals the necessity of a robust right of confrontation above and beyond a requirement of contemporaneous documentary evidence.\textsuperscript{274} As a result, Joshua’s suggestions are simultaneously too comprehensive in scope and too soft in recommendation. A superior solution would embrace *Crawford*’s requirement that testimonial statements may only be admitted if the declarant is available for cross-examination.\textsuperscript{275}

3. A Right of Confrontation in the Spirit of *Crawford*

One of the principal problems to arise from the 2002 Leniency Notice was the fact that procedures were implemented without preparing adequate procedural protections.\textsuperscript{276} While a number of commentators have recently advocated for increased procedural protections in EC competition hearings,\textsuperscript{277} to date the proposals have been infeasible. A more effective mechanism for securing procedural fairness for undertakings in EC competition hearings would allow cross-examination of adverse witnesses brought before the EC in the spirit of *Crawford v. Washington*.

Currently, Regulation 773/2004 states that “[t]he Hearing Officer may allow the parties to whom a statement of objections has been addressed, the complainants, other persons invited to the hearing, the Commission services and the authorities of the Member States to ask questions during the hearing.”\textsuperscript{278} An effective solution would add the following to that regulation:

Undertakings subject to a Statement of Objections shall have an opportunity to cross-examine any witness upon whose oral statements—transcribed or otherwise—the Commission intends to rely in proving an infringement of Article 101 of the Treaty. The Hearing Officer shall include a statement concerning the credibility of each testifying witness in his report.

\textsuperscript{274} *Crawford*, 541 U.S. at 68.
\textsuperscript{275} Since the oral hearing is the first opportunity for the prosecuted undertakings to hear the Commission’s witnesses live, the requirement of previous cross-examination of unavailable witnesses set forth in *Crawford* is inapplicable here. *Crawford*, 541 U.S. 36.
\textsuperscript{276} *International Cartels Roundtable*, supra note 19, at 100.
\textsuperscript{277} See discussion supra Parts IV.B.1-2.
\textsuperscript{278} Regulation No. 773/2004, supra note 177, art. 7, at 21.
This solution has a number of benefits. First, it avoids the necessity for formal criminalization of cartel conduct, which is currently impracticable within the EC. Additionally, as mentioned above, there is currently no desire among the member states of the EU, the General Court, and the ECJ to criminalize cartel conduct, and as such, formal criminalization is an ineffective means of achieving procedural fairness.

Although some may argue that a right of confrontation should only apply in criminal cases, and that therefore such a right is not necessary in EC hearings, the strong trend in cartel enforcement toward quasi-criminal investigations, an adversarial hearing model, and increasingly stiff financial penalties make the requirement that the hearings be criminal merely semantic. Moreover, despite the lack of imprisonment as a punishment for violation of Article 101, and the limitation of the American Confrontation Clause to “criminal prosecutions,” the imposition of crippling fines—affecting the lives and livelihoods of employees—even within a quasi-criminal framework provides a strong, prudential foundation for a right of confrontation in EC competition hearings.

Further, some might contend that the aforementioned proposal, which relies on an American case interpreting a jurisdiction-specific constitutional provision, represents an unjustified, anti-inquisitorial Anglocentrism. However, this challenge simultaneously ignores the reality of the English and American common law influence over the development of cartel enforcement in the EC, as well as the procedural trends toward a common law model that have taken shape within the EC itself. Therefore, even if the EC fails to fully criminalize cartel conduct, it should not be excused from extending key procedural rights that bear directly on truth-finding, especially when the procedures that already have been adopted so closely mirror those of common law systems. Lastly, irrespective of whether the EC disclaims the “criminal” nature of its cartel conduct.

279 Forrester, supra note 7, at 826.
280 Marsden & Whelan, supra note 186, at 35.
281 See supra Part III.
282 Baudenbacher, supra note 128, at 354.
283 U.S. CONST. amend. VI.
284 Sklansky, supra note 72.
285 Asimow & Dunlop, supra note 126, at 146.
286 See supra Part III.
287 See id.
enforcement regime, the system should, at base, be concerned with obtaining reliable, truthful evidence of the sort that only a robust right of confrontation can ensure. This is especially true where the credibility of its evidence-gathering techniques is in question.

Second, allowing cross-examination will provide an opportunity for targeted undertakings to directly test the credibility of the EC witnesses in the presence of a neutral hearing officer. In light of the *Stolt-Nielsen* decision, it is clear that the opportunity to cross-examine such witnesses is vital to a fair proceeding in the cartel enforcement context. Thus, while the proposal here may not affect the opinions of the DG Competition case team in its investigation, the Hearing Officer, charged with making credibility determinations, will need to make a decision as to the live credibility of each witness with respect to the statements collected in the leniency and investigation process. This will assuredly “improve the quality of the contradictory debate.” Indeed, it will also provide necessary procedural protections that will both improve the quality of the evidence garnered in EC competition proceedings and the fairness afforded those facing judgment.

It may be argued that placing the Hearing Officer in the role of a decision-maker with respect to credibility raises the same reliability concerns that the Supreme Court so forcefully criticized in *Crawford*, since the hearing officer will ultimately need to make a subjective determination about the credibility of the adverse witnesses. However, this challenge ignores the fact that the reliability test of *Roberts* was a threshold issue used to determine whether the out-of-court statements would be *admissible* as evidence, not whether the trier of fact would ultimately make a subjective judgment. Under the current proposal, the requirement of cross-examination means that any oral statement may only be used as a “roadmap” and not as evidence, unless the individual who made the statement is available for cross-examination. Moreover, because the Hearing Officer’s report is published in

---

288 See *supra* Part I.B.3.c.
289 See *supra* Part I.B.3.a.
291 Paulis, *supra* note 234, at 385-86.
293 *Id.*
the Official Journal of the European Union,294 credibility determinations will shine a light both on the quality of the evidence and the efficacy of the procedures. As a result, the proposed solution will incentivize the DG Competition case teams to produce credible witnesses and ensure testimonial accuracy.

Finally, the solution proposed here is incremental and legislative, avoiding the problems that attend comprehensive overhauls and judicially-mandated reforms.295 Creating a right of confrontation and cross-examination in the Crawford mold first solves the immediate problem of admitting untested and unsworn oral statements gathered in the investigatory stage,296 while, at the same time, recognizing that oral statements are vital in uncovering conspiracies and incentivizing undertakings to apply for leniency.297 Additionally, this proposal gives the accused an opportunity to challenge the credibility of witnesses before a neutral hearing officer, creating a disincentive for individuals and applicants for leniency to lie. Further, the proposal advocated here requires only a minor amendment to an existing regulation, incurring comparatively less institutional and monetary costs than competing “holistic” solutions, and representing a single-but-necessary step in an incremental approach toward a robust set of procedural rights. Although some may contend that the current proposal fails to go far enough, the strength of the proposal lies in the fact that it will allow the Commission and targeted undertakings to ease into durable procedural protections. In this way, the current proposal avoids the dangers of the initial effort to allow oral corporate statements brought about in the 2002 amendments to the Leniency Programme—which while an important advancement, lacked necessary foresight and procedural protections, providing too much room for abuse by the DG Competition and leniency applicants.

295 Forrester, supra note 7, at 840.
296 See supra Part IV.A.
297 van Barlingen & Barennes, supra note 149, at 9; Joshua, supra note 9, at 5-6.
V. CONCLUSION

The introduction of procedures for accepting oral statements in the leniency and investigatory contexts represented an important addition to EC competition procedure. However, the consequences of relying on oral statements as evidence against co-conspirators include both a shift toward the adversarial system of justice and a marked increase in the need for procedural protections to prevent abuse and to ensure testimonial accuracy. Centuries of common law have demonstrated that the best way to protect against overzealousness in law enforcement and dishonesty in testimonial statements by co-conspirators is through cross-examination. Therefore, in order to ensure procedural fairness in the EC, there must be a right for accused undertakings to confront and cross-examine those witnesses on whose oral statements the EC intends to rely as evidence.

Jameson J. Dempsey†

† J.D. Candidate, Brooklyn Law School, 2011; B.A., Boston College, 2005. Many thanks to the members of the Brooklyn Law Review for their input and tireless work editing my note. I would also like to extend my thanks to Professor Robert Pitler, Professor Yane Svetiev, and Professor Nelson Tebbe for their helpful comments throughout the writing process. Additional thanks to Kristen J. McAhren for introducing me to the topic and providing invaluable mentorship over the years. This note is dedicated to my family for their unconditional love and encouragement.