

1-1-2014

Fitting Science into Legal Contexts: Assessing Effects of Causes or Causes of Effects?

David Faigman

UC Hastings College of the Law, faigmand@uchastings.edu

Philip A. Dawid

Steven E. Feinberg

Follow this and additional works at: http://repository.uchastings.edu/faculty_scholarship

Recommended Citation

David Faigman, Philip A. Dawid, and Steven E. Feinberg, *Fitting Science into Legal Contexts: Assessing Effects of Causes or Causes of Effects?*, 43 *Soc. Methods & Res.* 359 (2014).

Available at: http://repository.uchastings.edu/faculty_scholarship/1307

This Article is brought to you for free and open access by UC Hastings Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of UC Hastings Scholarship Repository. For more information, please contact marcus@uchastings.edu.

Sociological Methods & Research

<http://smr.sagepub.com/>

Fitting Science Into Legal Contexts: Assessing Effects of Causes or Causes of Effects?

A. Philip Dawid, David L. Faigman and Stephen E. Fienberg

Sociological Methods & Research 2014 43: 359 originally published online 23
December 2013

DOI: 10.1177/0049124113515188

The online version of this article can be found at:

<http://smr.sagepub.com/content/43/3/359>

Published by:



<http://www.sagepublications.com>

Additional services and information for *Sociological Methods & Research* can be found at:

Email Alerts: <http://smr.sagepub.com/cgi/alerts>

Subscriptions: <http://smr.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Aug 14, 2014

[OnlineFirst Version of Record](#) - Dec 23, 2013

[What is This?](#)

Fitting Science Into Legal Contexts: Assessing Effects of Causes or Causes of Effects?

Sociological Methods & Research
2014, Vol. 43(3) 359-390
© The Author(s) 2013
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0049124113515188
smr.sagepub.com



**A. Philip Dawid¹, David L. Faigman²,
and Stephen E. Fienberg³**

Abstract

Law and science share many perspectives, but they also differ in important ways. While much of science is concerned with the effects of causes (EoC), relying upon evidence accumulated from randomized controlled experiments and observational studies, the problem of inferring the causes of effects (CoE) requires its own framing and possibly different data. Philosophers have written about the need to distinguish between the “EoC” and “the CoE” for hundreds of years, but their advice remains murky even today. The statistical literature is only of limited help here as well, focusing largely on the traditional problem of the “EoC.” Through a series of examples, we review the two concepts, how they are related, and how they differ. We provide an alternative framing of the “CoE” that differs substantially from that found in the bulk of the scientific literature, and in legal cases and

¹ Statistical Laboratory, Centre for Mathematical Sciences, Cambridge University, Cambridge, United Kingdom

² Hastings College of the Law, University of California, San Francisco, CA, USA

³ Department of Statistics, Heinz College, Machine Learning Department, and Cylab, Carnegie Mellon University, Pittsburgh, PA, USA

Corresponding Author:

A. Philip Dawid, Statistical Laboratory, Centre for Mathematical Sciences, Cambridge University, Cambridge CB3 0WB, United Kingdom.

Email: a.p.dawid@statslab.cam.ac.uk

commentary on them. Although in these few pages we cannot fully resolve this issue, we hope to begin to sketch a blueprint for a solution. In so doing, we consider how causation is framed by courts and thought about by philosophers and scientists. We also endeavor to examine how law and science might better align their approaches to causation so that, in particular, courts can take better advantage of scientific expertise.

Keywords

law, evidence, expert, causes of effects, effects of causes

Introduction

Law and science share many objects of concern, yet, as modes of understanding the world, they do not necessarily share the same perspective. This is particularly so in the context of scientific evidence, that is, when science is relied upon to answer factual disputes in litigation. In the courtroom, the object of concern is typically an individual, and the legal process is roughly tailored to determine, for instance, whether the civil defendant is responsible for the plaintiff's present condition or the criminal defendant is responsible for whatever he or she is accused of having done. In the laboratory or in science more broadly, in contrast, the object of concern is typically a population, and the scientific method is roughly tailored to determine, for instance, whether exposure to a substance increases or decreases the risk of a particular illness compared with nonexposure, or perhaps how one or another context affects behavior. Thus, one way to view this basic difference in perspective between science and law is that science studies individuals in order to make statements about populations, while the law considers populations in order to make statements about individuals. This essential difference in perspective between what scientists normally do and what the trial process is ordinarily about has profound implications for the law's use of scientific evidence, especially when it comes to matters of inferring causation.

How, and whether, to usefully employ scientific data 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, in which they distinguish routinely between "general causation" and "specific causation," a distinction that has a long history in philosophy and

science. Although, invariably, courts, like scientists, initially focus on the general question of causation, the ultimate issue in most trials concerns what caused the observed effect in a particular case. Indeed, many court cases begin with “the fact” that a party has a particular condition, for example, a disease or lower pay than “comparable” men, and the dispute involves the likely cause of that condition (Faigman 2010).

Courts generally understand, for example, that the question of whether benzene exposure significantly increases the risk of leukemia is fundamentally separate from the question of whether the plaintiff’s leukemia is attributable to benzene exposure. Knowing that benzene exposure causes leukemia does not, by itself, demonstrate that benzene exposure caused the plaintiff’s leukemia. But this produces a conundrum for the law and science connection. Science typically wishes to infer “the effects of causes” (EoC), through experiments and observational studies, but legal fact finders need to infer “the causes of effects” (CoE). Legal fact finders cannot reasonably infer from general data alone that a particular effect is attributable to a known cause, yet they are ultimately charged with exactly this responsibility—of determining at some level of certainty what caused a particular effect.

This question of how best to reason from group data to an individual case is pervasive in the courtroom use of science. Indeed, it is present in virtually all courtroom applications of scientific research. But it is rarely if ever cast as the question of inferring the CoE as opposed to the EoC. The courts have failed to bring a coherent strategy to this issue and are largely slapdash in their approach to this matter. Moreover, the distinction has not received the sustained attention of statisticians (but see Dawid 2000; Dawid 2011), even though it goes back to the writings of Mill (1873).

In this article, we seek to begin to remedy this situation. Ultimately, this subject is too big for the number of pages we have at our disposal. Nonetheless, we hope to offer some direction for how the challenge of using group data to decide individual cases might be approached, and in a more focused way alert both the courts and the scientists that the question of the CoE requires new ways of thinking about statistical and scientific evidence. The first section catalogs the several ways courts maneuver the intersection of law and science, where general data are employed to decide specific legal questions. The second section examines the different ways that scientists and philosophers think about causation, and considers whether these ideas might be of service to the courtroom use of science. The third section explores how science might be done so as better to inform legal policy, and the fourth section examines how different legal contexts make different demands on the

science. Finally, the fifth section seeks to frame the law's use of scientific experts, and considers implications for the intersection of law and science.

Making Inferences About Individual Cases in the Courtroom

In the courtroom, the question of reasoning from group data to a particular case occurs in just about every legal context in which scientific evidence is decided by the trier of fact (i.e., the jury or, if there is no jury, the judge). However, courts have not approached these cases with a uniform strategy, much less a strategy that is strongly informed by a sophisticated understanding of statistical inference, especially as it relates to issues of causation. Although it oversimplifies matters somewhat, it is fair to describe judicial practices in this area as comprising three strategies. The first limits the scope of expert testimony to statements about what the research indicates regarding the group, of which the specific case is believed to be an instance. The expert in this first category is limited to describing general research data, and the trier of fact is left to apply that information to the case at hand. Essentially, category 1 testimony permits expert opinion on the effects of the causes related to the issue in the present case, but leaves to the trier of fact the question whether the specific effect was a consequence of the putative cause. The second strategy permits expert testimony at both the general research level and its application to the particular case. The expert or experts in this category are allowed to testify regarding the general inferences permitted by the research, as well as to whether the present case is an instance of that general phenomenon. Essentially, category 2 testimony permits expert opinion on both the effects of certain putative causes and whether the cause of interest is responsible for the effect that is the subject of the legal dispute. The third strategy permits experts to reason directly from the (known or presumed) existence of the effect to its possible causes, including in particular the putative cause that is the subject of the case. This approach, which will seem most foreign to contemporary scientists, allows expert opinion that is based on the "clinical experience" of the expert. Essentially, category 3 testimony permits expert opinion on whether the effect was a consequence of the putative cause, despite the fact that there may be no data on the "EoC" in the respective area of concern.

Limiting Expert Testimony to the EoC

Although the ultimate issue in trials is usually case-specific, most evidence is general and merely helpful to the ultimate resolution of the case. In legal

parlance, such evidence is “circumstantial,” because some additional logical inference is needed for the fact finder to reason from the proof to the conclusion. The fact that, say, John, the defendant in a murder trial, stands to inherit a large sum of money as a consequence of the decedent’s death provides motive, but requires further deduction and other evidence. In contrast, testimony from an eyewitness that she saw John shoot the victim is considered “direct” proof, since no deduction is required. Of course, circumstantial evidence can have more or less weight than direct evidence. In the criminal case against John, for instance, all that might have been missing from the puzzle was motive. Alternatively, the eyewitness who testified she saw John shoot the decedent might be highly untrustworthy for a variety of reasons, including that she might have a motive to lie or has profoundly bad eyesight. In the process of legal proof, each piece of evidence must be considered on its own as regards its value in answering the ultimate legal question at issue. This makes clear one of the essential axioms of evidence law: that evidence can be admissible if it is probative regarding a fact in dispute.¹ Evidence need only to be a brick in the wall of proof, it need not be the wall. The same is true for scientific evidence.

Perhaps the paradigmatic example of scientific research whose use is limited by courts to informing triers of fact about possible general causes of an observed effect is testimony on eyewitness identification. A substantial body of research has identified a number of factors that interfere with eyewitness accuracy, with some affecting the initial identification and others the recall of the identification. Examples of the former include weapon focus and cross-racial identifications, while examples of the latter include confounding information received between identification and recall and the way the lineup is presented. Although the body of research on the causes of eyewitness inaccuracy is voluminous and robust (see Loftus [1979] 1996; Thompson et al. 1998; Wells and Olson 2003), most researchers would agree that it does not permit an expert to say with confidence whether a particular identification is accurate or not (Cutler and Wells 2009). To use a clinical term of art, the research cannot be used diagnostically to determine whether a particular eyewitness identification was accurate. Its relevance is limited to educating the jury about the limits of eyewitness identification.

Courts generally understand the probative value of eyewitness identification research as educational, in that it provides general contextual information to assist the trier of fact decide the case. In *United States v. Hines* (1999), for example, Judge Nancy Gertner emphasized that the virtue of this testimony is that the expert does not say that the eyewitness should be believed or disbelieved. “All that the expert does,” she explained, “is

provide the jury with more information with which the jury can then make a more informed decision” (p. 72). Professors John Monahan and Laurens Walker (1987) described this contextual information as “social framework” evidence, a term that has gained some currency in both the legal and scientific literatures (Faigman et al. 2013-2014).

According to Monahan and Walker (1987), social framework evidence has two components, the general social science findings (e.g., the fallibility of cross-racial identification) and its relevance and application to the particular case (e.g., whether the testifying eyewitness was inaccurate in his cross-racial identification). In the eyewitness area, as well as in many other areas, an expert could testify about the framework, but would not be permitted to opine about whether the particular case was an instance of the framework. The fact finders would thus be informed about the general research, but would have to make their own inferences about the case at hand.²

In the courtroom, most uses of scientific evidence involve the application of general principles to a specific case. Hence, even if the expert evidence is limited to the “EoC,” the ultimate legal issue to which the research is relevant concerns the case at hand. However, this is not always true. Sometimes a general proposition of science is itself the principal issue in a case. A good example of this is the case of *Entertainment Software Association v. Blagojevich* (2005), which involved an Illinois law restricting minors’ access to violent video games. 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” (p. 1073). 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” (p. 1073).

After extensively reviewing the psychological and neurological research advanced by the State, the court concluded that Illinois had failed to meet its considerable burden. 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” (p. 1074). 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” (p. 1074). The court permanently enjoined the Illinois law.

In *Entertainment Software*, the legally relevant inquiry was described at the group level, that is, whether a law limiting access to violent video games could be justified because such games have a deleterious impact on children as a whole. Thus, the “EoC” was the pivotal legal issue in dispute and scientific research could speak directly to this issue. The more typical courtroom situation, however, involves the relevance of scientific evidence to decide a case about a particular individual. In the context of violent videos, this might be whether a particular minor’s violent action can be attributed to his having played violent video games. Although the defense is unusual, defendants have on occasion argued insanity on the basis of video programming (Chananie 2007; Falk 1996). In *Zamora v. State* (1978), for example, “Zamora’s insanity defense was based on ‘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 claim, the defense 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 was allowed to testify. *Zamora* was convicted.

The *Zamora* case, in many respects, illustrates a situation that is exactly the opposite of what the model of scientific inference would presume. The court permitted the clinician to testify about the individual defendant, but not the psychologist who offered the research framework into which the individual purportedly fit. As in the case of expert testimony on eyewitness accuracy, the more enlightened choice would be the reverse, that is, to permit relevant and valid general research findings, but to allow expert opinion about the specific case only if it could be validly provided. The Alice in Wonderland approach of *Zamora* is considered in section I.C., subsequently. The more scientifically realistic problem of deciding whether to permit an expert to opine about how general research data fit a particular case is considered in the next section.

Permitting Expert Testimony on the Effects of Reputed Causes and the Cause of the Observed Effect

As described in the previous section, there are a number of areas in which the courts limit expertise to the general description of the EoC. The application of the insights provided by this general research to the case at hand is left to the fact finders to determine. At the same time, in many legal contexts, courts insist on expert proof at both levels of causation, general and specific. Indeed, especially in certain kinds of case, courts regularly expect experts to be able to provide testimony, both on the general effects of certain causes, and that the effect in the case at bar is attributable to a particular cause.³ The reasons *why* courts sometimes limit expert proof to the EoC, while at other times they insist on an opinion about the cause of a specific effect, are somewhat elusive. In most instances, an expert addressing the cause of a specific event will rely upon evidence of the EoC, whether from experimental evidence or some other form of data and analysis. Nonetheless, to a large extent, it appears that courts' expectations regarding the specificity of the proof are dictated by legal standards, rather than the ability of scientists to provide the necessary proof. The admission of expert testimony that some cause produced the observed effect depends primarily on the legal demand for such proof, rather than experts' ability to validly supply such proof.

The two broad areas in which courts routinely expect, and often require, proof of both general effects and specific causes are cases involving medical causation and forensic identification.⁴ While causation judgments in medicine and forensics are not alike scientifically, courts treat them as of a kind.⁵ As regards the former, courts have largely insisted that without adequate proof of both general causation and specific causation, the case fails. In the latter contexts, it appears that it was the forensic specialists who insisted on their ability to individualize, and courts have largely accepted this claim uncritically (Faigman, Monahan, and Slobogin forthcoming).

In the typical medical causation case, the plaintiff has the burden to prove, first, that some cause (i.e., drug, substance, or device) produces some deleterious effect and, second, that it did so in this case. Consider, for example, a claim that the plaintiff's leukemia was caused by his exposure to perchloroethylene (PCE), a chemical in dry-cleaning fluid (see *Magistrini v. One Hour Martinizing Dry Cleaning* 2002). Virtually all courts would require, first, credible scientific proof that PCE causes leukemia in populations exposed to it at dosages comparable to what the plaintiff was exposed to, and, second, that the plaintiff's leukemia was caused by it (see *In re Rezulin*

Productss Liability Litigation 2005). Neither component of this causation equation is simple or straightforward.

On the general causation side of the equation, credible scientific proof is likely to come in a variety of forms. In the case of PCE, a number of different research paradigms and a multitude of studies are likely to be cited both for and against the general connection claimed. These might include clinical observations, in vitro experiments, animal studies, and epidemiology. This work is likely to rely on uneven research designs, widely varying dosage levels, and a range of possible statistical interpretations. Nonetheless, the basic question of whether the available research converges sufficiently to permit the conclusion that PCE causes leukemia is one that is familiar to scientists.

Significantly less familiar to scientists is the second component of the causation equation, assessing whether the alleged cause, in fact, produced the observed effect. In medical causation cases, courts have largely settled on a single means by which experts are permitted to reason from evidence of general causation to proof of specific causation. This method is referred to as “differential etiology.”⁶ Differential etiology is a strategy of formal logic devised by lawyers and not a medical or scientific term. In short, differential etiology requires a testifying expert, first, to rule in the putative cause and then, second, to rule out other possible causes of the effect in issue. This method has little scientific grounding and, indeed, is as much art as science.

To be done right, differential etiology would involve a multitude of factors, few of which are easily quantified or yet well defined. Without attempting to provide a comprehensive guide to how differential etiology might be done validly, we note that a testifying expert would have to attend to a multiplicity of factors.

As an initial starting point, 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 expert must initially consider the strength of the relationship between *X* and *Y* as a general matter. For example, both second-hand smoke and firsthand smoke are associated with lung cancer, but the strength of the relationship is generally 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. 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 scientific findings that are supported by adequate research—requires two abilities, neither of which is clearly within most scientists' skill sets. The first, and perhaps less problematic, is that of forensic investigator. Almost no matter what the empirical relationship alleged, exposure or dosage levels will be relevant to the diagnosis. The first principle of toxicology is that "the dose makes the poison," since any substance in sufficient quantities could injure or kill. 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 of cause must integrate the surfeit of information concerning the multitude of factors that make up the general model, combine this with the case history information known or suspected about the individual, and offer an opinion with some level of confidence that substance *X* was the likely cause of condition *Y*. In practice, this opinion is usually stated as follows: "Within a reasonable degree of medical certainty, it is my opinion that *X* caused [a particular case of] *Y*." This expression has uncertain empirical meaning and is mainly a mantra repeated by experts for purposes of legal decision makers, who similarly have little 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. Such statements are cloaked in the guise of science, but are little more than educated guesses. Indeed, experts' case-specific conclusions appear to be based largely on an admixture of an unknown combination of knowledge of the subject, experience over the years, commitment to the client or cause, intuition, and blind faith.

Despite the lack of a scientific basis for differential etiology, in the medical causation context the perceived need to have an expert connect general

research to the case at hand is, in many respects, understandable. Unlike expert evidence that indicates that cross-racial identifications are less accurate on average than same-race identifications, expert evidence that PCE doubles the risk of leukemia will have little meaning to the average juror. Expert testimony on eyewitness accuracy educates the fact finder to the possible limitations associated with a powerful item of evidence. The expert opinion helps jurors reach a conclusion about the individual case, because they can then integrate this information into the totality of the other evidence available. In medical causation cases, however, jurors need more. Knowing that PCE doubles the risk is helpful, certainly, but without some guidance regarding how to apply this information to the case at hand, jurors are unlikely to know what to do with this information. At least, this appears to be the judgment reached by courts, since they uniformly require credible testimony on both general and specific causation.

In civil cases involving medical causation, therefore, the imperative to admit testimony about whether the specific effect is an instance of some general effects is more the product of the demands of the law than the power of the science. On the criminal side of the judicial docket, a parallel phenomenon occurs, but without the legal need. Non-DNA forensic identification experts regularly seek to testify regarding the specifics of particular cases. Forensic experts are infamous for claiming the ability to “individualize” trace evidence (Saks and Koehler 2008). This ability, they claim, permits them to say that the latent fingerprint, bullet, or shoeprint came from the finger, gun, or shoe in question—to the exclusion of all other fingers, guns, or shoes in the world. Although the claimed reasoning process that would permit forensic individualization is not identical to that of differential etiology, it is similarly without scientific basis. Moreover, unlike the legal demands of medical causation cases, there is no need in criminal cases for experts to opine about the specific case. Indeed, any such imperative runs in the other direction.

In the non-DNA forensic identification sciences, such as fingerprints, firearms, handwriting, bitemarks, and so forth, the basic task is one of pattern recognition. Although the different areas present varying challenges, the essential charge for each of these specialties can be described similarly. An unknown finger, gun, shoe, or set of teeth left a pattern at the scene of a crime that is similar to or consistent with a known finger, gun, shoe, or set of teeth associated with a suspect or the defendant. From a scientific standpoint, the initial question concerns the frequency of the pattern. A sneaker print associated with a size 11 men’s running shoe will have a higher random match probability than a similar shoe in size 14. Forensic examiners,

however, dwell little on such notions of statistical probability, and search for allegedly unique wear patterns. Their objective is to say that the unknown pattern left at the scene of the crime indisputably came from the known pattern associated with the defendant (*United States v. Havvard* 2000).

Like the case-specific testimony offered in medical causation cases, somewhere between the general framework evidence that the case at hand *could* be an instance of the general phenomenon, a miracle occurs and some expert offers the opinion that the case at hand *is* an instance of that general phenomenon. In the most scientifically sophisticated form of forensic identification science, DNA profiling, no such divine intervention is needed. Instead, because the science has advanced adequately to permit substantial confidence in the frequency of finding the DNA evidence in a random sample of the population, DNA evidence is effectively presented at the general or framework level. Although the statistics might be overwhelming, such as a random match probability of 1 in 10 billion, the expert does not state that the sample came from an identified person to the exclusion of everyone else in the world.⁷

If non-DNA forensic identification experts were limited to what the science can bear, they would be allowed to testify only about the general data—assuming that data are available to support this testimony. In fact, in many forensic specialties, no such general data exist (Saks and Faigman 2008). In effect, in these areas, forensic examiners reason from the supposed causes of the particular effect to the effects of pertinent causes. This is a phenomenon that occurs even more regularly in medical causation cases, in which experts sometimes seek to testify that the temporal association between a cause and an effect provides proof of the general effect of that cause. This approach turns the scientific method on its head. Yet, many courts make special allowance for such testimony, a practice that is the focus of the next section.

Permitting Experts to Reason From the Existence of the Effect to Possible Causes

The idea that courts might allow experts to rely on the fact that a particular effect immediately followed a purported cause, to support their conclusion that that cause produced that effect, is one that surely would make most scientists' teeth hurt. Indeed, it violates that age-old caution, *post hoc ergo propter hoc* (“after this, therefore because of this”). Yet, many courts, especially state courts, permit such inferences under what is essentially an exception to their general rule for scientific testimony. This rule, referred to as the

“opinion rule,” allows experts to rule in a suspected cause of an illness or condition on the basis of their experience with the subject. In practice, the opinion rule allows experts to use differential etiology to rule in the disputed cause, despite the absence of research studies that would support that hypothesis. Not surprisingly, the opinion rule is justified primarily on the basis of legal precepts, rather than any scientific claims.

The opinion rule is primarily a function of the hoary *Frye* test (*Frye v. United States* 1923), which requires that a scientific technique or opinion should have achieved general acceptance in its particular field before it can be admitted in court. The *Frye* test itself is generally perceived as a fairly conservative test that limits the amount of expert evidence that appears in court. The opinion rule operates as an exception to the demand of general acceptance. In contrast to *Frye*, the predominant rule applied in courts today comes from the Supreme Court’s decision in *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993; Faigman and Monahan 2005). The *Daubert* Court held that judges had a responsibility as gatekeepers to ensure that the methods and principles underlying proffered expert opinion were more likely than not reliable and valid. The basic distinction between *Frye* and *Daubert* is one of perspective. *Frye* requires judges to ask the respective fields from which the science comes whether it is “generally accepted,” while *Daubert* puts the onus on judges to inspect the scientific foundation for the proffered testimony.

The opinion rule is primarily a feature of *Frye* jurisdictions, though only a minority of them subscribe to it. *Daubert* jurisdictions, possibly because of the greater scientific sophistication demanded of them, have largely eschewed the practice of permitting post hoc reasoning, at least in medical causation cases.⁸ A good example of the fairly sophisticated federal approach to this issue is *Hendrix v. Evenflo Co.* (2010). In *Hendrix*, the plaintiff claimed that the traumatic brain injury suffered by a newborn in a traffic accident caused the victim’s autism spectrum disorder. The trial court excluded the testimony and the Eleventh Circuit Court of Appeals affirmed. The appellate court noted, initially, that general research does not support the hypothesis that traumatic brain injury is a cause of autism (p. 1196). Thus, the expert had failed to rule in the operative cause of the condition. The expert, however, had sought to effectively rule in trauma as the cause of autism by a combination of the temporal proximity between the trauma and the condition and, through differential etiology, ruling out all other known causes. The Eleventh Circuit rejected this method as scientifically indefensible. “[A] mere temporal relationship between an event and a patient’s disease or symptoms does not allow an expert to place that event on a list of

possible causes of the disease or symptoms” (p. 1197). Indeed, the court added that anecdotal experience alone was not enough to rule in the putative cause. “Case studies and clinical experience, used alone and not merely to bolster other evidence, are also insufficient to show general causation” (p. 1197).

In stark contrast to *Hendrix* is the case of *Hood v. Matrixx Initiatives, Inc.* (2010). In *Hood*, a Florida trial court had excluded the plaintiff’s expert, who offered to testify that the plaintiff’s use of a nasal gel had caused him to lose all sense of smell. However, the intermediate appellate court reversed, holding that the expert’s testimony qualified as pure opinion and thus the *Frye* rule did not apply. According to the appellate court, “‘pure opinion’ testimony does not have to meet *Frye* because this type of testimony is based on the expert’s personal experience and training” (p. 1173). The court explained that the Florida opinion rule was mandated by the Florida Supreme Court’s decision in *Marsh v. Valyou* (2007). Under *Marsh*, expert opinion “developed from inductive reasoning based on the experts’ own experience, observation, or research” is not subject to *Frye* analysis (p. 1173). *Frye* is restricted, under this view, to “when an expert witness reaches a conclusion by deduction, from applying new and novel scientific principle, formula, or procedure developed by others” (p. 1173).

The *Marsh* approach, and that of the opinion rule more generally, is fundamentally contrary to basic precepts of the scientific method, if not logic itself. After all, if the complained-of substance or product has not been shown to be a scientifically plausible cause of the condition, how can an expert opine that it was the cause? Justice Cantero raised this concern in dissent in *Marsh*, arguing that testimony about specific causation must necessarily depend on proof of general causation. He observed as follows:

The majority’s holding that an opinion about specific causation need not pass the *Frye* test, even where the underlying theory of general causation is not accepted, in effect renders specific causation testimony always admissible as the “pure opinion” of the expert. . . . Permitting an expert to testify that X caused Y in a specific case without requiring the general acceptance of the theory that X can *ever* cause Y expands the “pure opinion” exception to the point where it swallows the rule. (pp. 562-63)

The basis for the opinion rule is not mere ignorance alone, but rather a judgment regarding the purposes of the civil jury system. The *Hood* court readily recognized that several federal courts, applying *Daubert*, had excluded the very expert to be admitted here. But, the court stated, it felt “compelled to

find that [his] opinion is admissible in Florida under *Marsh*” (p. 1175). The *Hood* court explained that *Marsh* held that

it is unnecessary for a plaintiff to conclusively demonstrate a causal link or to identify the “precise etiology” of the medical condition allegedly caused by the substance or predicate event. Accordingly, *Marsh* presents a “battle-of-the-experts” approach to the admissibility of expert testimony, designed to prevent trial judges from usurping “the jury’s role in evaluating the credibility of experts and choosing between legitimate but conflicting scientific views.” (p. 1175, quoting *Marsh*, p. 549)

The *Hood* court’s candor that the opinion rule is about legal process, not scientific method, is also in accord with the California approach. California, like Florida, historically permitted an exception to its *Frye*-oriented rule for pure opinion testimony.⁹ California based its opinion rule on perceptions of judicial process rather than the nature—or validity—of the expert evidence. According to the California Supreme Court, the primary concern with scientific opinion is the possibly overwhelming influence it might have on jurors (*People v. Kelly* 1976). The court observed that “[l]ay jurors tend to give considerable weight to ‘scientific’ evidence when presented by ‘experts’ with impressive credentials” (p. 1245). The *Frye* threshold, therefore, establishes a barrier in order to protect against juror credulity. Nonscientific expert opinion, according to the California high court, does not present a similar danger: “When a witness gives his personal opinion on the stand—even if he qualifies as an expert—the jurors may temper their acceptance of his testimony with a healthy skepticism born of their knowledge that all human beings are fallible” (*People v. McDonald* 1984).

It is not obvious that juror credulity is functionally different as between expert scientific opinion based on deduction and expert opinion based on pure experience (Faigman 2008). It is well to recall, however, the Florida Supreme Court’s statement in *Marsh* that the opinion rule is legitimate only so long as it gives to the jury the opportunity of “choosing between legitimate but conflicting scientific views.” The scientific legitimacy of reasoning from a specific effect to general causes is considered in the following sections.

Different Ways to Think About Causation

While statisticians have sometimes recognized the distinction between the EoC and the CoE, their focus has been largely on the former. For example, in a widely cited article on causality in science, Holland (1986) makes this

Table 1. Joint Probabilities.

	$Y = 0$	$Y = 1$
$X = 0$	p_{00}	p_{01}
$X = 1$	p_{10}	p_{11}

distinction, but goes on to focus on models for the EoC and the importance of randomization in inferring these. Indeed, Holland (1986, 1993) and others who cite him usually focus on measuring the “average causal effect” as described in the work of Rubin (1974), Rubin (1978), and Splawa-Neyman ([1923a] 1990), which is estimable precisely in situations involving randomization.

In his discussion of the history of causal thinking, Holland notes how it appears in various guises in the writings of Aristotle, Lock, Hume, and especially John Stuart Mill. He quotes Mill as follows:

... we have not yet proved that antecedent to be the cause until we have reversed the process and produced the effect by means of that antecedent artificially, and if, when we do so, the effect follows, the induction is complete . . . (p. 252)

Although this focuses more on the EoC, Mill goes on to note:

Induction is mainly a process for finding the causes of effects: and . . . in the more perfect of the sciences, we ascend, by generalization from particulars, to the tendencies of causes considered singly, and then reason downward from those separate tendencies, to the effect of the same causes when combined.

... as a general rule, the EoC are far more accessible to our study than the CoE . . .

While statisticians do not often distinguish between the EoC and the CoE, they do distinguish between prospective and retrospective studies, as do epidemiologists. In situations involving categorical data, for example, when there are two putative causes ($X = 0$ and $X = 1$) and two possible outcomes ($Y = 0$ and $Y = 1$), statistical analysts focus on the probabilities in a 2×2 contingency table of Table 1, where p_{00} , p_{01} , p_{10} , and p_{11} are nonnegative “cell probabilities” that sum to 1. Then the association between X and Y is often measured in terms of the odds ratio (OR):

$$\text{OR} = \frac{p_{00}/p_{01}}{p_{10}/p_{11}} = \frac{p_{00}/p_{10}}{p_{01}/p_{11}} = \frac{p_{00}p_{11}}{p_{01}p_{10}}.$$

The important property of the OR is that it is symmetric as between X and Y ; in particular, it can be estimated both from prospective studies where we “control” or condition on X and then examine Y , and from retrospective studies, where we control or condition on Y and then observe X . This follows from the crucial property of the OR that it is unchanged by rescaling of rows and columns (Bishop, Fienberg, and Holland 1975; Edwards 1963; Fienberg 1980). An extension of this equivalence relates to the validity of computing logistic regression coefficients, with data from a retrospective study, just as if the data had been collected prospectively—then we are essentially looking at adjusted ORs (Farewell 1979).

The OR differs from a criterion widely used by epidemiologists and the courts: the *relative risk* (or *risk ratio*), RR. Whereas the OR compares the relative *odds* of the outcome Y for the two values of X , the *relative risk* compares the probabilities, P_1 and P_0 , of the outcome for the two values of X . When P_1 and P_0 are small, OR approximates RR. However, when P_1 and P_0 are close to .5, the OR is typically much larger than RR.

The RR, or some form of adjusted RR, is often, and often quite reasonably, used as a measure of the effect (on Y) of a cause (X); however, its application to assessing the cause of an effect is more problematic. Some courts require that RR exceed 2 before invoking the legal criterion of “more probable than not.”¹⁰ This comes from the epidemiological literature on “assigned shares” (Lagakos and Mosteller 2006) or attributable risk (Greenland and Robins 1988). However, as we will see below this is not correct, at any rate if taken as implying that a value of RR below 2 does not support a claim of causation at the civil standard of proof.

EoC Versus CoE

It might be thought, from the relevance of the OR, in a 2×2 table to both prospective and retrospective studies, that it could be used to elide the distinction between the EoC and the CoE. However, this is not the case, since it turns out that it is the relative risk RR, and not the OR, that is required for assessing CoE; and RR does not share OR’s insensitivity to the direction of inference—just as the general equivalence of prospective and retrospective inference fails when either X or Y is continuous, as was clear in the work of Galton who, in the nineteenth century, showed that the regression of X on Y is not obtainable from the regression of Y on X .

Table 2. Comparative Experiment.

	No Recovery	Recovery	Total
Chalk	88	12	100
Aspirin	70	30	100

To illustrate the relationships and differences between “EoC” and “CoE,” we consider the following pair of questions (Dawid 2011):

EoC: I have a headache. I am wondering whether to take aspirin. Will that cause my headache to disappear?

CoE: I had a headache and took aspirin. My headache went away. Was that caused by the aspirin?

We attempt to address these two queries using the data in the 2×2 contingency table of Table 2, constructed from a hypothetical double-blind randomized experiment involving 200 individuals, with 100 given aspirin tablets, and 100 chalk tablets (the control). The patients take their assigned tablets the next time they get a headache, and record how long it is until the headache has gone. We interpret “recovery” as (say) “headache disappears within 30 minutes.”

The OR is 3.14, so we have some evidence that taking an aspirin causes headaches to go away. Adjusting for covariates might sharpen the inferences, by removing additional sources of variation, but the randomized experimental results in Table 3 above are about as good as it gets for inferring the effect (on recovery) of the cause (taking aspirin).

Suppose, however, we are interested in the CoE, as we are in the legal context. So we consider a situation where someone did take aspirin ($A = 1$), and the actual response was recovery ($R = 1$). We might now imagine two “potential responses”: R_0 , indicating recovery (or not) if the aspirin is *not* taken, and R_1 , indicating recovery if it *is* taken. We have observed $R_1 = 1$, but have not observed—indeed, we now have no way of observing— R_0 . To address the issue of whether taking the aspirin caused the observed recovery, we might ask: What is the probability that the (necessarily unobserved) potential response R_0 , which would have been observed had I *not* taken aspirin ($A = 0$), would have been different ($R_0 = 0$) from that actually observed ($R_1 = 1$)?

It would obviously be nice if we could estimate this probability from data, such as that in Table 3, but this is not straightforward. From these results, we can only fill in the margins of Table 3, as shown. Since we can never observe

Table 3. Potential Responses.

R_1	R_0		Totals
	0	1	
0	$88 - x$	$x - 18$	70
1	x	$30 - x$	30
Totals	88	12	100

both responses R_0 and R_1 on the same individual, as a matter of logic we cannot observe x , the number of individuals having $R_0 = 0$ and $R_1 = 1$. However, as soon as we hypothesize a value for x , the remaining figures in the table are determined, as shown.

We can then estimate the probability of causation (PC) as:

$$PC = \Pr(R_0 = 0 | R_1 = 1) = x/30.$$

Now because all the entries in Table 3 must be nonnegative, we know that $x \geq 18$ and thus $PC \geq .6$. The value of PC is thus greater than the civil standard of proof, .5, so we have evidence to support the conclusion that the aspirin was the cause of the recovery. But note that, even if we had perfect information from the experiment, with no uncertainty, we can say little more about PC (unless we have some additional relevant scientific information).

Applying the above argument with completely general frequencies, we can relate the PC to the causal relative risk $\Pr(R_1 = 1)/\Pr(R_0 = 1)$ —which will be the same as RR so long as we derive this quantity from a properly performed and analyzed experimental study. We then find:

$$PC \geq \{\Pr(R_1 = 1) - \Pr(R_0 = 1)\} / \Pr(R_1 = 1) = 1 - 1/RR.$$

In particular, $RR > 2$ implies that $PC > 1/2$. This gives some support to the courts' inclination to require relative risks to exceed 2 in order to meet the civil standard of proof. Because there is only an inequality relationship between PC and RR, however, finding that RR falls short of 2 does not necessarily imply that $PC < 1/2$.

We caution that, depending on the context, there may be other valid ways to make inferences about the CoE than through the measure PC. Also, if we have additional information, for example, on genetic effects moderating the effects of treatment, we may be able to refine further our inferences about PC (Dawid 2010). But in any event, even if we start with the best possible

information (perfect experimental results) about the EoC, and use all relevant auxiliary information, we need to apply subtle logic to make inferences about the CoE (which will still, necessarily, remain imprecisely determined).

Dawid, Musio, and Fienberg (2013) provide a more general treatment of the general framework for bounding PC and for extensions to it involving uncertainty in the data on the potential outcomes.

Specifics of CoE in a Legal Context

Discrimination Litigation

The language of civil rights laws dating back to the 1960s dealing with discrimination has a strong causal flavor. Title VII of the 1964 Civil Rights Act prohibits an employment practice that affects an individual's employment "because of such individual's race . . ." Similar language arises in sex discrimination law. We observe that a woman is not hired for a job at company *X*. Essentially the law asks: What would have happened if the applicant for a job had been a man rather than a woman but nothing else had changed? Obviously, we cannot observe this counterfactual situation and compare it with what actually happened, and thus the law has traditionally settled for observational data and statistical methods such as regression for measuring pay disparities and logistic regression for disparities in hiring and promotion. Expert witnesses using such methods employ the "sex coefficient" to measure discrimination. We observe the outcomes for those who applied to company *X* or were in the company *X* workforce and then we ask whether sex was a cause, and so we consider a regression of pay on sex and other permissible explanatory variables that measure qualifications, market factors, and so on. A recent report from the National Research Council notes:

To draw inferences from running regressions on observational data, substantial prior knowledge about the mechanisms that generated the data must be used to support the necessary assumptions. Studies vary substantially in the degree to which the necessary assumptions are adequately justified. (Blank, Dabady, and Citro 2004)

What typically goes unsaid in such settings is that we are really interested in the cause of the observed effects, as opposed to the effects of the cause (sex), which is what the regression actually measures. We can't simply reverse the regression (even if that were a well-defined notion), and, unlike what we saw for the OR, the "sex coefficient" is not the same going forward and backward. In the 1980s, these issues were at the center of several discrimination

lawsuits, and a number of statisticians wrote about the role of reverse regression for measuring discrimination. Conway and Roberts (1983) and Dempster (1988) and the accompanying discussion of these articles provide a readable summary of the technical debate on this topic, but that tended to focus on the “fairness” of qualifications, rather than the contrast between CoE and EoC—though this is implicit in the language of the academic debate. Academic interest in the question of what regression to use and how to capture the question before the courts then waned (but see Greiner 2008), and courts continued to see a battle of experts presenting alternative regression analyses of a more traditional sort. Rarely is there any discussion of how these analyses relate to the actual company processes for hiring and deciding on pay and promotions.

Wal-Mart v. Dukes (2011) was to be the largest class-action lawsuit in the history of the United States. The Plaintiffs were current or former employees of Wal-Mart. They sought judgment against Wal-Mart for injunctive and declaratory relief, punitive damages, and back pay, on behalf of themselves and a nationwide class of some 1.5 million female employees, because of what they alleged to be Wal-Mart’s discrimination against women in violation of Title VII of the Civil Rights Act of 1964.

Under applicable legal standards for certifying the class, the Plaintiffs were obligated to demonstrate that “there are questions of law or fact common to the class” (p. 2549). In the case at hand, this required that class members “have suffered the same injury” (p. 2551). At the class certification hearing in federal district court, Plaintiffs’ sociological expert witness testified regarding his “social framework analysis” of Wal-Mart’s “culture” and personnel practices, and concluded that the company was “vulnerable” to gender discrimination. The reasoning here was from the general—that of Wal-Mart’s “strong corporate culture”—to the specific—that Wal-Mart discriminated against its women employees as a consequence (Mitchell, Monahan, and Walker 2011a, 2011b).

The Plaintiff’s economic expert reported on regression analyses for compensation, region by region with dummy variables for stores. In these analyses, the sex coefficient was highly statistically significant, but this was not surprising given the large number of employees included in the analyses. The Defendant’s expert also reported on regression analyses with additional explanatory variables, but at the store level and even department within store level. In her analyses, the sex coefficient was largely insignificant or was not well identified.

In the lower courts, the Plaintiffs’ experts’ theories prevailed. The Supreme Court, however, found the evidence presented by the Plaintiffs’

experts inadequate. Indeed, the Court concluded that the putative class was too broad and lacked commonality, even if the Plaintiffs' experts' testimony was accepted at face value. More generally, as regards the sociological evidence, the Court went on to note that "we can safely disregard what he has to say. It is worlds away from 'significant proof' that Wal-Mart 'operated under a general policy of discrimination.'" Justice Ginsburg, concurring in part and dissenting in part, did not mention the Plaintiffs' sociological evidence, but she did find the regression analyses credible. She and the other dissenters from this part of the Court's opinion believed that the majority had incorrectly interpreted the rules of class certification, and would have remanded the case for further consideration.

Epidemiological Evidence of Causation: The Case of Thimerisol and Autism

A 1998 *Lancet* publication linked the measles, mumps, and rubella (MMR) vaccine to the emergence of autism in 12 young children within 1–14 days after receiving the vaccine (Wakefield et al. 1998). There were no comparable controls and no adjustments for covariates, let alone a randomized controlled clinical trial. Subsequently, attention was focused on thimerisol, a preservative widely used in multiple use vials of the MMR vaccine, as a potential causal agent, because it contained ethyl mercury and methyl mercury, which had been linked to a variety of neurological disorders. An elaborate case–control study was ultimately carried out by the Centers for Disease Control, and this and other epidemiological studies were reviewed by a series of Committees at the Institute of Medicine. They reported that "the body of epidemiological evidence favors rejection of a causal relationship between thimerisol-containing vaccines and autism. The committee further finds that potential biological mechanisms for vaccine-induced autism that have been generated to date are theoretical only" (Immunization Safety Committee 2004). Nonetheless, many lawsuits followed, filed by parents of autistic children against the vaccine manufacturers and the U.S. Government. According to Moreno (2009),

By January 2009, 5,535 cases alleging that vaccines cause autism had been filed against the Department of Health and Human Services. On February 12, 2009, the first three autism "test cases" were decided by the United States Court of Federal Claims Office of Special Masters under the National Vaccine Injury Compensation Program. (the "Federal Vaccine Court")

The three test cases were heard by separate special masters.¹¹ Typical of the voluminous expert witness testimony on behalf of the petitioners was that of a toxicologist, who opined that autism could be caused by a disorder that prevents children from effectively eliminating mercury. This assertion was countered by a defense expert who claimed that the available evidence did not support such a link and that “it is inappropriate to generalize . . . from one form of mercury to another” (Moreno 2009, p. 1525). In the background was the epidemiologic evidence in the 1998 *Lancet* article.

Although the cases were heard separately the special masters concurred in the following language:

First, the specific/individual causation claims were rejected because the medical evidence proffered by “the petitioners . . . failed to demonstrate that her[his] vaccinations played *any role at all* in causing those [autism-related] problems.” Second, the general causation claims were rejected because “petitioners . . . *failed* to demonstrate that thimerosal-containing vaccines . . . or that the MMR vaccine *can* contribute to causing . . . autism” (Moreno 2009:1514)

Much has been written about the outcome of these cases, which have virtually all been dismissed at this time (Brady 2009; Moreno 2009; Offit 2007; Sugarman 2007). The original research article that triggered much of the controversy has since been withdrawn by the editors because of allegations of fraud and the lead author of the article has been censured by the British Medical Association and struck off the U.K. medical register (Hawkes 2011).

Framing the Law’s Use of Scientific Experts

There remains a yawning disconnect between how the law defines expert proof and the ability of science validly to supply such proof. To some extent, this divide is merely a product of the law’s unrealistic expectations about the power of science. Applied science—which encompasses all of the science that the law cares about—is limited by the methods, both research and statistical, that are available to it in particular contexts. The law, in contrast, constructs legal doctrines, either with utter disregard for what scientists might be able to say about the facts relevant to those doctrines or in light of a folk understanding of the science that might bear on those relevant facts. In either case, a more clear-eyed view of the fundamental distinction between identifying, on one hand, the EoC and, on the other hand, the CoE, will do much to lessen the divide between these two disciplines.

Law and science are entirely separate institutions, with distinct approaches to identifying what is important to consider or study, to the methodologies brought to bear on those subjects deemed relevant, and to the objectives or goals of the entire enterprise. At the same time, these two venerable institutions intersect in a multitude of ways. When they do meet, however, each does so on its own terms, or, that is, from its own perspective. From the scientist's standpoint, law very often suggests hypotheses of interest or permits the application of hypotheses to new and salient contexts. From the law's standpoint, science very often supplies the brute facts integral to legal decision making or policy formation or creates new matters that legal doctrine must address. In short, when law and science intersect, the two institutions continue to maintain their separate methods and objectives, with each eyeing the other for its own purposes. The law and science intersection can only be described from each of their relative perspectives (Faigman 1999).

While law and science are, and inevitably will remain, separate, the issue we address in this article lies dead center at their point of intersection. The issue of reasoning from group data to individual decision (also referred to as G2i)—that is, having proof of both the EoC and the CoE—presents a key challenge regarding the translation of scientific findings into legal outcomes. It is perhaps not surprising that this translational issue has not shone more brightly on the radar screens of both lawyers and scientists. As we illustrated in the previous sections, for scientists, reasoning about the CoE is in tension with conventional methods of studying populations and inferring causation; and for lawyers, such reasoning appears solvable by recourse to conventional logic (i.e., differential etiology) or folk science. Thus, neither scientists nor lawyers have delved deeply into the subject. But given their fundamentally different identities and sensibilities, each must approach the matter of the CoE in its own way.

From the perspective of science, the challenge of reasoning from group data to individual decisions is, at first glance, a statistical one. But even brief reflection on the subject soon reveals the deep complexity of the subject. This complexity operates on both the front end of research design and on the back end in applying the findings to particular cases. On the front end, a scientist's ability to increase the specificity of findings can depend on innumerable factors, including the sampling method used, data collected, research design employed, and statistical tests applied. On the back end, it might be possible to define and regularize the concept of differential etiology more than has so far been done. Given the fairly amorphous state of this methodology, at least insofar as the courts make use of it, there is considerable room for improvement.

From the perspective of the law, the challenge of reasoning from group data to individuals is endemic to the whole enterprise. In the trial setting, the issue is not so much the inevitable need to individualize group data, but the question of when the law should, and indeed whether it must, turn that function over to experts. Under virtually all evidence codes, expert evidence must be helpful to the fact finder. Although this principle does not mean that the subject must be entirely outside the ken of jurors, it does contemplate that experts will provide jurors with more information, and give them more assistance, than they would otherwise have. Added to this basic helpfulness requirement for expert evidence is the basic axiom that evidence need only be probative, it need not prove the case: Evidence is admissible if it is a brick in the wall, it does not have to be the wall.

When experts present evidence regarding the individual case, they potentially run afoul of these two principles of evidence law. Specifically, if science cannot validly speak to individual cases—that is, if it cannot identify the cause of an identified effect—then their testimony is not helpful. Moreover, the inclination of expert witnesses to resolve the ultimate issue in the particular case, most evident in medical causation and non-DNA forensic identification cases, manifests the perception that scientific evidence needs to be more than merely a brick in the wall of proof.

At the start of this article, we noted that our subject was too big for the number of pages at our disposal. And indeed it is, for any redress of the challenges we have attempted to identify here will take the concerted effort of both scientists and legal scholars.

It is clear to us that scientists could do much more in studying the matter of reasoning about the CoE. This is, at bottom, a topic infused with methodological and statistical issues, and is highly relevant not only to sundry legal contexts but also to medicine and public health. At the same time, it is clear to us that the law must similarly take a hard look at the state of the science of inferring CoE and consider the necessity of admitting this evidence without better safeguards. The law's analysis of this issue will likely have to be done context by context, if not case by case. As we have noted, looking at the usual scientific question of the EoC precedes reasoning about the CoE, and there may be limited scientific evidence to bring to bear on the latter question. Nonetheless, in most medical causation cases, the demands of the subject may lead courts to require testimony about the CoE. The framing of expert testimony regarding discriminatory behavior in employment may take a very different form. In either case, however, the determination of what courts expect from expert testimony should be explicit, and should depend on both the judge's and the jury's need for the help and the expert's ability validly to

provide that help. In other contexts, such as forensic identification in criminal cases, this assessment might come out differently.

Authors' Note

The contents reflect the views of the authors and do not necessarily represent the official views of either the John D. and Catherine T. MacArthur Foundation or the MacArthur Foundation Research Network on Law and Neuroscience (www.lawneuro.org).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Preparation of this article was supported, in part, by a grant from the John D. and Catherine T. MacArthur Foundation.

Notes

1. Federal Rule of Evidence 401 defines “relevant evidence” as “evidence having any tendency to make 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” (Federal Rule of Evidence, 2013). However, this rule needs to be interpreted carefully. Thus, in a case of identification through a DNA match between a trace taken at the crime scene and a sample from a suspect, either one of these DNA profiles taken singly is, according to a strict interpretation of the rule, of no evidential value, and so to be excluded as not relevant; however, taken together, the profiles become highly probative.
2. The general findings presented in evidence would concern the prospective accuracy of the process of identifying some given individual (of given race); the trier of fact, however, must assess the “retrospective” accuracy of a given identification. The required “transposition of the conditional” requires the (formal or informal) application of Bayesian reasoning—failing which it is prey to serious misinterpretations, such as the “prosecutor’s fallacy.” See, for example, Thompson and Shumann (1987). Nonetheless, courts typically do not provide any guidance on this delicate feat of reasoning.
3. Very often parties use different experts to provide proof of the general and proof of the specific.
4. A third area in which courts routinely expect proof of both general and specific causation is, in some ways, a combination of the two discussed in the text. In the area of behavioral forensics, the courts often insist on both general and specific

proof, though, as in medicine and forensic identification, the general research is often weak or absent altogether. Experts in this area, usually psychiatrists and psychologists, testify about a wide range of subjects, such as posttraumatic stress disorder, competency, predictions of violence, mental abnormality, battered woman syndrome, and so forth. Many of the observations made in the text about medical causation and forensic identification are fully applicable to behavioral forensics.

5. Courts generally do not invest substantial effort in explaining why they sometimes allow, and even sometimes insist that expert proof be forthcoming in regard to both the general EoC and the specific cause of an effect. In the routine practice of adjudication, however, medical and forensic experts routinely testify regarding whether this particular case is an instance of some general phenomenon. Courts treat the issue as if they partake of the same logic. They do not. Forensic identification is typically concerned with assessing uncertainty about an unknown past event, such as “did the suspect leave this DNA trace?” This is essentially an exercise in Bayesian inference and might be termed “backcasting.” In contrast, medical causation is an exercise in assessing a causal connection between known events (e.g., exposure and disease) and might be termed “attribution.”
6. Some courts have confused differential etiology with “differential diagnosis.” Whereas differential diagnosis concerns identification of the illness, differential etiology concerns assessment of the cause of the illness. Diagnosis precedes etiology, and in many medical settings it is the end of the inquiry. For medical purposes, knowing that the individual has leukemia is the operative concern. In the law, the diagnosis of illness is a necessary prerequisite, but the cause of the illness is typically the fact that is in dispute.
7. Some further reasoning toward this end, such as the presentation and explanation of the “likelihood ratio”—the valid measure of the strength of the match evidence—may however be permitted.
8. Federal courts take a lax approach to forensic science cases and, more generally, permit some forms of expert testimony on the basis of experience alone. Courts are especially permissive, in this regard, concerning police officers. Other experiential experts might include auto mechanics, real estate appraisers, and harbor pilots. The ostensible reason for permitting such experience-based expertise is courts’ belief that it is sufficient to support valid testimony. It is worth pointing out that the primary reason underlying the opinion rule is not validity, but that experience-based opinion does not present the dangers of overwhelming the jury with the aura of scientific exactitude.
9. Although it remains too soon to determine, California may be moving away from using the opinion rule, as evidenced by the recent decision in *Sargon Enterprises v. Univ. of So. Cal.* (Cal. 2012). In *Sargon*, the California Supreme Court held

that trial judges must operate as gatekeepers—a term borrowed from *Daubert*—to ensure that the material the expert relies upon “actually supports the expert’s reasoning” at 1252. This gatekeeping function applies to all experts.

10. In *In re Silicone Gel Breast Implants Products Liability Litigation* (C.D.Cal. 2004), the court observed:

When statistical analyses or probabilistic results of epidemiological studies are offered to prove specific causation, however, under California law those analyses must show a relative risk greater than 2.0 to be “useful” to the jury . . . [A] relative risk of 2.0 implies a 50% probability that the agent at issue was responsible for a particular individual’s disease. This means that a relative risk that is greater than 2.0 permits the conclusion that the agent was more likely than not responsible for a particular individual’s disease.

See also *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (9th Cir. 1995) [*Daubert II*] (The court noted that a relative risk below 2.0 [and above 1.0] may suggest a causal relationship, “but it actually tends to disprove legal causation, as it shows that Bendectin does not double the likelihood of birth defects.”).

11. *Cedillo v. Secretary of Health and Human Services* (Fed. Cl. February 12, 2009), *Hazlehurst v. Secretary of Department of Health and Human Services* (Fed. Cl. February, 12 2009), and *Snyder ex rel. Snyder v. Secretary of Department of Health and Human Services* (Fed. Cl. February 12, 2009).

References

- Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland. 1975. *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press (reprinted by Springer, 2007).
- Blank, R. M., M. Dabady, and C. F. Citro, eds. 2004. *Measuring Racial Discrimination*. Atlanta, Georgia: Panel on Methods for Assessing Discrimination, National Research Council, National Academy Press.
- Brady, M. T. 2009. “Commentary: Court Decisions Should Allay Parents’ Vaccine Fears.” *American Academy of Pediatrics News* 30:1.
- Cedillo v. Secretary of Health and Human Services*, No. 98-916V, 2009 WL 331968 (Fed. Cl. February 12, 2009).
- Chanania, J. 2007. “Violent Videogames, Crime, and the Law: Looking for Proof of a Causal Connection.” *Developmental Mental Health Law* 26:27-44.
- Conway, D. A. and H. V. Roberts. 1983. “Reverse Regression, Fairness, and Employment Discrimination.” *Journal of Business & Economic Statistics* 1:75-85.
- Cutler, B. and G. Wells. 2009. “Expert Testimony Regarding Eyewitness Identification.” Pp. 100-123 in *Psychological Science in the Courtroom: Controversies and Consensus*, edited by J. Skeem, K. Douglas, and S. Lilienfeld. New York: The Guilford Press.

- Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).
- Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 43 F.3d 1311, 1321 (9th Cir. 1995).
- Dawid, A. P. 2000. "Causal Inference without Counterfactuals (with Discussion)." *Journal of the American Statistical Association* 95:407-28.
- Dawid, A. P. 2011. "The Role of Scientific and Statistical Evidence in Assessing Causality." Pp. 133-47 in *Perspectives on Causation*, edited by R. Goldberg. Oxford, England: Hart Publishing.
- Dawid, A. P., M. Musio, and S. E. Fienberg. 2013. "From Statistical Evidence to Evidence of Causality." Submitted for publication.
- Dempster, A. P. 1988. "Employment Discrimination and Statistical Science (with Discussion)." *Statistical Science* 3:149-75.
- Edwards, A. W. F. 1963. "The Measure of Association in a 2x2 Table." *Journal of the Royal Statistical Society Series A* 126:109-14.
- Entertainment Software Association v. Blagojevich*, 404 F.Supp.2d 1051 (N.D.Ill. 2005).
- Faigman, D. L. 1999. *Legal Alchemy: The Use and Misuse of Science in the Law*. New York: W. H. Freeman.
- Faigman, D. L. 2008. "Admissibility Regimes: The "Opinion Rule" and Other Oddities and Exceptions to Scientific Evidence, the Scientific Revolution, and Common Sense." *Southwestern University Law Review* 36:699-722.
- Faigman, D. L. 2010. "Evidentiary Incommensurability: A Preliminary Exploration of the Problem of Reasoning from General Scientific Data to Individualized Legal Decision-making." *Brooklyn Law Review* 75:1115-36.
- Faigman, D. L., J. A. Blumenthal, E. K. Cheng, J. Mnookin, E. A. Murphy, and J. Sanders. 2013-2014. *Modern Scientific Evidence: The Law and Science of Expert Testimony*. Eagan, MN: Thomson/West.
- Faigman, D. L. and J. Monahan. 2005. "Psychological Evidence at the Dawn of the Law's Scientific Age." *Annual Review of Psychology* 55:631-59.
- Faigman, D. L., J. Monahan, and C. Slobogin. forthcoming. "Group to Individual Inference (G2i) in Scientific Expert Testimony." *University of Chicago Law Review* 81.
- Falk, P. J. 1996. "Novel Theories of Criminal Defense Based upon the Toxicity of the Social Environment: Urban Psychosis, Television Intoxication, and Black Rage." *North Carolina Law Review* 74:731-811.
- Farewell, V. 1979. "Some Results on the Estimation of Logistic Models Based on Retrospective Data." *Biometrika* 66:27-32.
- Federal Rules of Evidence. 2013. *Prepared by Daniel J. Capra*. Minneapolis, MN: West Publishing.
- Fienberg, S. E. 1980. *The Analysis of Cross-classified Categorical Data*. 2nd ed. Cambridge, MA: MIT Press (reprinted by Springer, 2007).

- Frye v. United States*, 293 F. 1013 (App.D.C. 1923).
- Greenland, S. and J. M. Robins. 1988. "Conceptual Problems in the Definition and Interpretation of Attributable Fractions." *American Journal of Epidemiology* 128:1185-97.
- Greiner, D. J. 2008. "Causal Inference in Civil Rights Litigation." *Harvard Law Review* 122:533-98.
- Hawkes, N. 2011. "College Investigates Whether Wakefield was Guilty of Scientific Fraud." *British Medical Journal* 342. doi:10.1136/bmj.d2010
- Hazlehurst v. Secretary of Department of Health and Human Services*, No. 03-654V, 2009 WL 332306 331968 (Fed. Cl. February 12, 2009).
- Hendrix v. Evenflo Co., Inc.*, 609 F.3d 1183 (11th Cir. 2010).
- Holland, P. W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81:945-60.
- Holland, P. W. 1993. "What Comes First, Cause or Effect?" Pp. 273-82 in *A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues*, edited by G. Keren and G. Lewis. Hillsdale, NJ: Lawrence Erlbaum.
- Hood v. Matrixx Initiatives, Inc.*, 50 So.2d 1166 (Dist.Ct.App.Fla. 2010).
- Immunization Safety Committee. 2004. *Immunization Safety Review: Vaccines and Autism*. Washington, DC: Institute of Medicine, National Research Council, National Academies Press, Executive Summary.
- In re Rezulin Products Liability Litigation*, 369 F.Supp.2d 398, 401-02 (S.D.N.Y. 2005).
- In re Silocone Gel Breast Implants Products Liability Litigation*, 318 F.Supp.2d 879, 893 (C.D.Cal. 2004).
- Lagakos, S. W. and F. Mosteller. 2006. "Assigned Shares in Compensation for Radiation-related Cancers." *Risk Analysis* 6:345-57.
- Loftus, E. F. [1979] 1996. *Eyewitness Testimony. Revised Edition*. Cambridge, MA: Harvard University Press.
- Magistrini v. One Hour Martinizing Dry Cleaning*, 180 F.Supp.2d 584 (D.N.J. 2002).
- Marsh v. Valyou*, 977 So.2d 543 (Fla. 2007).
- Mill, J. S. 1873. *The Collected Works of John Stuart Mill, Volume VII—A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation* (Books I-III), edited by John M. Robson, Introduction by R. F. McRae. Toronto, Canada: University of Toronto Press, London: Routledge and Kegan Paul, 1974).
- Mitchell, G., J. Monahan, and L. Walker. 2011a. "Case-specific Sociological Inference: Metanorms for Expert Opinions." *Sociological Methods & Research* 40: 668-80.
- Mitchell, G., J. Monahan, and L. Walker. 2011b. "The ASA's Missed Opportunity to Promote Sound Science in Court." *Sociological Methods & Research* 40 605-20.

- Monahan, J. and L. Walker. 1987. "Social Frameworks: A New Use of Social Science in Law." *Virginia Law Review* 73:559-97.
- Moreno, J. A. 2009. "It's Just a Shot Away: MMR Vaccines and Autism and the End of the Daubertista Revolution." *William Mitchell Law Review* 35:1511-41.
- Offit, P. A. 2007. "Thimerosal and Vaccines—A Cautionary Tale." *New England Journal of Medicine* 357:1278-79.
- People v. Kelly*, 549 P.2d 1240, 1248-51 (Cal. 1976).
- People v. McDonald*, 690 P.2d 709, 724 (Cal. 1984).
- Rubin, D. B. 1974. "Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies." *Journal of Educational Psychology* 66:688-701.
- Rubin, D. B. 1978. "Bayesian Inference for Causal Effects. The role of Randomization." *Annals of Statistics* 6:34-58.
- Saks, M. J. and D. L. Faigman. 2008. "Failed Forensics: How Forensic Science Lost Its Way and How It Might Yet Find it." *Annual Review Law and Social Science* 4: 149-71.
- Saks, M. J. and J. J. Koehler. 2008. "The Individualization Fallacy in Forensic Science Evidence." *Vanderbilt Law Review* 61:199-219.
- Sargon Enterprises v. Univ. of So. Cal.*, 288 P.3d 1237 (Cal. 2012).
- Snyder ex rel. Snyder v. Secretary of Department of Health and Human Services*, No. 01-162V WL 332044 331968 (Fed. Cl. February 12, 2009).
- Splawa-Neyman, J. [1923a] 1990. "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9." *Statistical Science* 5: 465-72. Translated and edited by D. M. Dabrowska and T. P. Speed. from the Polish original, which appeared in *Roczniki Nauk Rolniczye*, Tom X (1923): 1-51 (*Annals of Agricultural Sciences*).
- Sugarman, S. D. 2007. "Cases in Vaccine Court—Legal Battles over Vaccines and Autism." *New England Journal of Medicine* 357:1275-77.
- Thompson, C. P., D. Herrmann, J. D. Read, D. Bruce, D. G. Payne, and M. P. Togli. 1998. *Eyewitness Memory: Theoretical and Applied Perspective*. Mahwah, NJ: Lawrence Erlbaum.
- Thompson, E. L. and E. L. Shumann. 1987. "Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy." *Law and Human Behavior* 2:167.
- United States v. Havvard*, 117 F.Supp.2d 848, 852 (S.D. Ind. 2000).
- United States v. Hines*, 55 F.Supp.2d 62 (D.Mass. 1999).
- Wakefield, A. J., S. H. Murch, A. Anthony, J. Linnell, D. M. Casson, M. Malik, M. Berelowitz, A. P. Dhillon, M. A. Thomson, P. Harvey, A. Valentine, S. E. Davies, and J. A. Walker-Smith. 1998. "Ileal-lymphoid-nodular Hyperplasia, Non-specific Colitis, and Pervasive Developmental Disorder in Children [retracted]." *Lancet* 351:637-41.

Wal-Mart v. Dukes, 131 S.Ct. 2451, 2011 WL 2437013 (2011).

Wells, G. and E. Olson. 2003. "Eyewitness Testimony." *Annual Review of Psychology* 54:277-96.

Zamora v. State, 361 So.2d 776, 779 (Fla.Dist.Ct.App. 1978).

Author Biographies

A. Philip Dawid is Emeritus Professor of statistics at the University of Cambridge. He has been awarded the Royal Statistical Society's Guy Medal in Silver for his wide-ranging research, which includes a long-standing interest in the interpretation of evidence and causal reasoning. His recently edited volumes include "Simplicity, Complexity and Modelling," "Evidence, Inference and Enquiry," and "Causality: Statistical Perspectives and Applications," as well as the book of the Darwin College Cambridge Lecture Series on "Beauty."

David L. Faigman is John F. Digardi Distinguished Professor of law at the University of California, Hastings College of the Law. He also holds an appointment in the School of Medicine, Department of Psychiatry, at the University of California San Francisco and is codirector of the UCSF-UC Hastings Consortium on Law, Science and Health Policy. His scholarship primarily concerns the integration of scientific research into legal decision making. His most recent book is *Constitutional Fictions: A Unified Theory of Constitutional Facts*, and he is a coauthor/coeditor of the five volume treatise *Modern Scientific Evidence: The Law and Science of Expert Testimony*.

Stephen E. Fienberg is Maurice Falk University Professor of Statistics and Social Science at Carnegie Mellon University and director of the Living Analytics Research Centre, a joint activity with Singapore Management University. Best known for his work on categorical data analysis, he has long-standing interests in the use of statistics in legal contexts and causal inference more broadly. He is a member of the National Academy of Sciences and a fellow of the American Academy of Arts and Sciences, and the Royal Statistical Society.