Scientific and Expert Evidence

ASPEN CASEBOOK SERIES

Scientific and Expert Evidence

Third Edition

John M. Conley
William Rand Kenan Jr. Professor of Law
University of North Carolina School of Law

Jane Campbell Moriarty
Carol Los Mansmann Chair in Faculty Scholarship and
Professor of Law
Duquesne University School of Law

ASPEN
PUBLISHING

To contact Customer Service, e-mail customer.service@aspenpublishing.com,
call 1-800-950-5259, or mail correspondence to:

Aspen Publishing
Attn: Order Department
PO Box 990
Frederick, MD 21705

Printed in the United States of America.

1 2 3 4 5 6 7 8 9 0

ISBN 978-1-4548-9792-7

Library of Congress Cataloging-in-Publication Data

Names: Conley, John M., author. | Moriarty, Jane Campbell, author.
Title: Scientific and expert evidence / John M. Conley, William Rand Kenan Jr. Professor of Law, University of North Carolina Schoolof Law; Jane Campbell Moriarty, Carol los Mansmann Chair in Faculty Scholarship and Professor of Law, Duquesne University School of Law.
Description: Third edition. | Frederick, MD: Aspen Publishing, [2021] | Series: Aspen casebook series | Includes index. | Summary: "Cases and materials for elective course on scientific and expert evidence for law students"—Provided by publisher.
Identifiers: LCCN 2020029234 (print) | LCCN 2020029235 (ebook) | ISBN 9781454897927 (hardback) | ISBN 9781543819625 (ebook)
Subjects: LCSH: Evidence, Expert—United States.
Classification: LCC KF8961 .C665 2020 (print) | LCC KF8961 (ebook) | DDC 347.73/67—dc23
LC record available at https://lccn.loc.gov/2020029234
LC ebook record available at https://lccn.loc.gov/2020029235

About Aspen Publishing

Aspen Publishing is a leading provider of educational content and digital learning solutions to law schools in the U.S. and around the world. Aspen provides best-in-class solutions for legal education through authoritative textbooks, written by renowned authors, and breakthrough products such as Connected eBooks, Connected Quizzing, and PracticePerfect.

The Aspen Casebook Series (famously known among law faculty and students as the "red and black" casebooks) encompasses hundreds of highly regarded textbooks in more than eighty disciplines, from large enrollment courses, such as Torts and Contracts to emerging electives such as Sustainability and the Law of Policing. Study aids such as the *Examples & Explanations* and the *Emanuel Law Outlines* series, both highly popular collections, help law students master complex subject matter.

Major products, programs, and initiatives include:

- **Connected eBooks** are enhanced digital textbooks and study aids that come with a suite of online content and learning tools designed to maximize student success. Designed in collaboration with hundreds of faculty and students, the Connected eBook is a significant leap forward in the legal education learning tools available to students.
- **Connected Quizzing** is an easy-to-use formative assessment tool that tests law students' understanding and provides timely feedback to improve learning outcomes. Delivered through CasebookConnect.com, the learning platform already used by students to access their Aspen casebooks, Connected Quizzing is simple to implement and integrates seamlessly with law school course curricula.
- **PracticePerfect** is a visually engaging, interactive study aid to explain commonly encountered legal doctrines through easy-to-understand animated videos, illustrative examples, and numerous practice questions. Developed by a team of experts, PracticePerfect is the ideal study companion for today's law students.
- The **Aspen Learning Library** enables law schools to provide their students with access to the most popular study aids on the market across all of their courses. Available through an annual subscription, the online library consists of study aids in e-book, audio, and video formats with full text search, note-taking, and highlighting capabilities.
- Aspen's **Digital Bookshelf** is an institutional-level online education bookshelf, consolidating everything students and professors need to ensure success. This program ensures that every student has access to affordable course materials from day one.
- **Leading Edge** is a community centered on thinking differently about legal education and putting those thoughts into actionable strategies. At the core of the program is the Leading Edge Conference, an annual gathering of legal education thought leaders looking to pool ideas and identify promising directions of exploration.

To Paula

J.M.C.

To Tom

J.C.M.

Summary of Contents

Contents

Preface

Preface to the First Edition

We bring to this book a mixture of academic and practical experience with scientific evidence, a blend that we hope is reflected in the book itself. For each of us, the book represents a logical next step in a lengthy relationship with the topic. Both of us spent substantial apprenticeships as big-firm litigators, during which we dealt extensively with expert witnesses, both friendly and hostile. We both have also dealt with the judiciary on these issues, Conley as director of a "judging science" program and a teacher of law and social science at the University of Virginia's Graduate Program for Judges, and Moriarty as clerk to a state supreme court justice with a special interest in questions of law and science. And both of us have written on law and science, Moriarty primarily on psychological and "forensic science" issues, and Conley on anthropology and statistics.

At about the time that *Daubert* brought new urgency to the topic, each of us began teaching a course on scientific evidence. Although we were not in consultation at the time, each of us concluded that, to be an intelligent and critical consumer of scientific evidence, a lawyer needs a grounding both in the relevant legal doctrines and in the basic scientific principles that underlie various types of evidence—not that a lawyer needs to function as a scientist (neither of us fits that description), but it is not sufficient to treat the "scientific" part of scientific evidence as a black box to be managed by the experts. We have both believed from the outset that mastery of a relatively few overarching scientific concepts and processes can greatly enhance a lawyer's effectiveness. Accordingly, as we assembled, tested, and revised our own *ad hoc* teaching materials, we particularly sought cases and other sources that deal with scientific issues in some depth, in a way that is both accurate and straightforward.

The completion of the *Daubert* trilogy; the burgeoning academic literature about the theoretical, practical, and policy implications of the trilogy; the revision of Federal Rule of Evidence 702; and, most importantly, the proliferation of significant cases that seem to turn on the admissibility and sufficiency of scientific evidence acted in concert to persuade us that the time had come to turn our *ad hoc* teaching materials into a casebook. In simplest terms, scientific evidence had become part of the basic literacy of every courtroom lawyer, whether civil or criminal, and that reality seemed to call for a course with a "real" book.

We hope that law students and their teachers will find this to be a challenging yet non-intimidating introduction to the scientific techniques that regularly enter the courtroom and the evidentiary principles that govern their use. If we, who majored in Latin literature (Conley) and philosophy (Moriarty), have been able to achieve elemental literacy, then so can you. We hope that this book will make your journeys a good deal less painful than ours have been.

John M. Conley
Jane Campbell Moriarty

May 2007

Preface to the Third Edition

The purpose of the third edition is twofold: first, to reflect significant new developments in both law and science; and second, to improve the clarity with which we conveyed the scientific material in the original edition. Toward the first end, almost every chapter contains new cases and scientific materials. However, because most teachers are likely to be using the book in a three-hour course, we have tried to make its profile even slimmer by deleting superseded entries. In addition, we have done some significant reorganizing that we hope will make the book more efficient for teachers and students. Most importantly, this edition contains entirely redesigned chapters covering, respectively, developments in Opinion Evidence, including new cases exploring the complexity and boundaries of expert evidence that are suitable for student projects, and Social Science, Behavioral Science, and Neuroscience, with new cases and commentary.

In pursuing enhanced clarity, we have, with the help of some faculty users of the book, gone through all our scientific expositions with a close and critical eye, looking for places where we could explain things more simply, which sometimes also means at greater length. For example, in response to numerous comments from teachers, we have provided an expanded explanation of basic statistical concepts, with additional examples and illustrations. In a few instances, we have deleted scientific material from the first edition that experience has shown to be too esoteric and not critical to a law school course in scientific evidence. Finally, we have substantially revised and expanded the Teacher's Manual.

We hope that the new edition will prove even more user-friendly to teachers and students, and we hope that you will pass along your comments.

John M. Conley
Jane Campbell Moriarty

August 2020

Acknowledgments

I owe thanks to three generations of students at the University of North Carolina School of Law for their outstanding research assistance: Rhiannon D'Agostin, for the first edition; Ryan Gladden, for the second; and Brian Champion, Zan Newkirk, and Janet Witchger, for the third

My longtime friend David Peterson also deserves a particular expression of gratitude. A forensic statistician whose work has been relied on by the Supreme Court in multiple cases, David has been my mentor (and frequent collaborator) on all things statistical since I began teaching in 1983. He generously read and commented on the statistical components of this book.

Jake Barnes, Scott Baker, and Richard Myers were kind enough to read and comment on portions of the manuscript within their respective areas of expertise, as were the anonymous reviewers solicited by Aspen for the original edition, and the faculty users of the book who made suggestions for the second and third. My students (including many judges) over the years have also been astute reviewers of the materials as they have developed.

And above all, thanks to my wife, Paula, for everything.

J.M.C.

From 1997 to 2011, I had the pleasure of working at the University of Akron School of Law, which is where I created materials that became a course in Scientific and Expert Evidence. Neither the first nor second edition of this book would have been written without the input, encouragement, and assistance of the administration, my wonderful colleagues, students, and many dear friends at Akron Law. To all of you, my deeply felt affection and appreciation. Particular thanks to Tracy Thomas, Elizabeth Reilly, Will Huhn, J. Dean Carro, Marty Belsky, and Bill Jordan for the friendship, mentoring, and encouragement on this project.

Three terrific research assistants, now all *summa cum laude* graduates of Akron Law School helped with every phase of the second edition: Marisa Main, Erick Rigby, and Niki Ford. You all did remarkably fine work, and I am very grateful.

For the third edition, thanks to the support from Duquesne University School of Law and my Duquesne Law students, Stephen Reddy (2020), Kallie Crawford (2021), and Kyle Baicker-McKee (2021). You did great work under the most stressful conditions. You're going to be fine lawyers.

Michael Saks, John Conley, Will Huhn, Simon Cole, Bill Thompson, Gail Mason, Michael Perlin, Adina Schwarz, Mindy Mechanic, Maureen O'Connor, Michael Risinger, Mark Brodin, David Faigman, Daniel Langleben, Jennifer Mnookin, Joelle Moreno, Dale Nance, Michael Pardo, Paul Giannelli, Daniel Medwed, Georgio Ganis, David Harris, Erica Beecher-Monas, Hank Greely, Jay Koehler, Martha Farah, Myrna Raeder, Peter Rosenfeld, John Meixner, Erin Murphy, Ed Imwinkelried, Francis X. Shen, Andrea Roth, Jennifer Bard, Deborah Denno, and Emily Murphy have helped shape my thoughts on various aspects of science and evidence over the last several years. I'm grateful to each of you. Thanks to the anonymous reviewers solicited by Aspen who provided many helpful suggestions both for the first and second editions.

The most important person to thank is my husband Tom Lahman, who makes everything possible. You're the best part of every day.

J.C.M.

We would also like to acknowledge the authors, publishers, and copyright holders of the following publications for permission to reproduce excerpts herein:

Angell, Marcia, Editorial: Do Breast Implants Cause Systemic Disease?— Science in the Courtroom, 330 New England J.Med. 1784 (1994), reprinted with permission of Massachusetts Medical Society and New England Journal of Medicine. Copyright 1994 Massachusetts Medical Society. All rights reserved.

Jones, Owen D., et al., Brain Imaging for Legal Thinkers: A Guide for the Perplexed, 2009 Stan. Tech. L. Rev. 5, with permission from Owen D. Jones.

Matshes, Evan W., M.D. and Sam W. Andrews, Sam W., M.D. The Autopsy as a "Dying" Art, 42 MAR Champion 34 (2018)Lentini, John, Fire, Arsons & Explosions: Introductory Discussion of the Science, reprinted from Modern Scientific Evidence: The Law and Science of Expert Testimony, 2019-2020, with permission of Thomson Reuters.

Line drawing of a tire that was published in the United States Supreme Court case, Kumho Tire Co., Ltd. v. Carmichael (1999). Art courtesy of Kumho Tire and reprinted with permission by Mike Mavrigian.

Moriarty, Jane C., and Michael J. Saks, Forensic Science: Grand Goals, Tragic Flaws and Judicial Gatekeeping, published in Judges' Journal, Volume 44, No. 4, Fall 2005. Copyright © 2005 by the American Bar Association. Reprinted with permission by the General Counsel of the American Bar Association.

National Research Council (U.S.), Committee on Identifying the Needs of the Forensic Science Community, Strengthening Forensic Science in the United States: A Path Forward, reprinted with permission from the National Academies Press, Copyright © 2009, National Academy of Sciences.

Peterson, David W., Analyst Takes Fresh Look at Pay Discrimination (2006), reprinted with permission of the author.

Thomson, William C., and Dan E. Krane, DNA in the Courtroom, in Jane Campbell Moriarty, Psychological and Scientific Evidence in Criminal Trials (2003), reprinted with permission of Thomson Reuters.

Scientific and Expert Evidence

Chapter 1

Introduction to Science and the Legal Process

"Scientific evidence" is a far more complicated phrase than it might appear at first glance. To lawyers, it usually means "expert testimony based on scientific knowledge and inquiry that is offered in litigation"—in other words, *evidence*, in the legal sense, whose source is *scientific*. A lawyer's concern is typically with the legal questions of whether scientific evidence will be admitted in court and how much weight it will be given, as well as with the practical question of its persuasive impact.

To a scientist, however, the phrase and its constituent words are likely to mean something quite different—although exactly what is not clear. To start with, "the word *evidence* is used much more loosely in science than in the law. The law has precise rules of evidence that govern what is admissible and what isn't. In science the word merely seems to mean something less than 'proof.'" David Goodstein, *How Science Works*, in *Reference Manual on Scientific Evidence* 51 (Federal Judicial Center, 3d ed. 2011).* And the question of what counts as "scientific" has been debated by philosophers of science for

*Throughout this book, we make frequent reference to the Federal Judicial Center's *Reference Manual on Scientific Evidence*. The Center is the research and training arm of the federal judiciary, and it prepared the *Reference Manual* as a guide for federal judges who have to deal with various kinds of scientific evidence. As we write, the current edition of the *Reference Manual* is the third edition, published in 2011. For more information and a free downloadable version, readers can go to www.fjc.gov.—Eds.

centuries, with no resolution in sight. So science lacks a precise, universally accepted definition of either "evidence" or "scientific."

Why should lawyers care about these philosophical and linguistic issues? Until 1993, most American courts, both state and federal, applied the "*Frye* test" and admitted purportedly scientific evidence if the subject matter had "gained general acceptance in the particular field to which it belongs." Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923). A trial judge had only to look to the expert's community of practice and decide, aided by testimony and/or published works, the straightforward factual question of whether the work being offered had achieved broad acceptance. Whether it was deserving of such acceptance was irrelevant. Now, however, that simple determination is not enough. In the federal courts and the majority of the states, trial judges must now evaluate whether "scientific" evidence is really scientific, *according to the standards of science*. Under the "*Daubert* trilogy" of United States Supreme Court cases—Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993); General Electric Co. v. Joiner, 522 U.S. 136 (1997); and Kumho Tire Co. v. Carmichael, 526 U.S. 135 (1999)—and the state decisions that have adopted their standards, trial judges are required to make an independent determination of whether proffered scientific testimony is "ground[ed] in the methods and procedures of science." *Daubert*, 509 U.S. at 590. Judges must therefore have some understanding of what "the methods and procedures of science" are.

The legal requirements imposed by the *Frye* and *Daubert* tests are dealt with in great detail in Chapter 2. The purpose of this chapter is to introduce, in general terms, the problems that can arise when judges must apply scientific standards that they may not adequately understand, and about which scientists themselves may not agree. The cases and materials that follow will address such issues as the judge's duty to identify scientific knowledge, how scientists define what is "scientific," how judges think science works, and what can happen when the two are in conflict.

I. THE JUDGE AS ARBITER OF SCIENCE

Daubert v. Merrell Dow Pharmaceuticals, Inc.

509 U.S. 579 (1993)

Justice Blackmun delivered the opinion of the Court.

In this case we are called upon to determine the standard for admitting expert scientific testimony in a federal trial.

I

Petitioners Jason Daubert and Eric Schuller are minor children born with serious birth defects. They and their parents sued respondent in California state court, alleging that the birth defects had been caused by the mothers' ingestion of Bendectin, a prescription antinausea drug marketed by respondent. Respondent removed the suits to federal court on diversity grounds.

After extensive discovery, respondent moved for summary judgment, contending that Bendectin does not cause birth defects in humans and that petitioners would be unable to come forward with any admissible evidence that it does. [The district court held that the petitioners' expert testimony was inadmissible under the *Frye* general acceptance standard and the Ninth Circuit affirmed. The Supreme Court reversed, holding that Federal Rule of Evidence 702 requires independent judicial scrutiny of the reliability and relevance of scientific evidence.]

. . . [U]nder the [Federal] Rules [of Evidence] the trial judge must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable.

The primary locus of this obligation is Rule 702, which clearly contemplates some degree of regulation of the subjects and theories about which an expert may testify. "*If scientific*, technical, or other specialized *knowledge will assist the trier of fact* to understand the evidence or to determine a fact in issue" an expert "may testify *thereto*." (Emphasis added.) The subject of an expert's testimony must be "scientific . . . knowledge."[8] The adjective "scientific" implies a grounding in the methods and procedures of science. Similarly, the word "knowledge" connotes more than subjective belief or unsupported speculation. The term "applies to any body of known facts or to any body of ideas inferred from such facts or accepted as truths on good grounds." Webster's Third New International Dictionary 1252 (1986). Of course, it would be unreasonable to conclude that the subject of scientific testimony must be "known" to a certainty; arguably, there are no certainties in science. See, *e.g.*, Brief for Nicolaas Bloembergen et al. as *Amici Curiae* 9 ("Indeed, scientists do not assert that they know what is immutably 'true'—they are committed to searching for new, temporary, theories to explain, as best they can, phenomena"); Brief for American Association for the Advancement of Science et al. as *Amici Curiae* 7–8 ("Science is not an encyclopedic body of knowledge about the universe. Instead, it represents a *process* for proposing

8. Rule 702 also applies to "technical, or other specialized knowledge." Our discussion is limited to the scientific context because that is the nature of the expertise offered here. [The Court developed a more flexible standard for nonscientific experts in *Kumho Tire*, the third case in the *Daubert* trilogy. See Chapter 2.]

and refining theoretical explanations about the world that are subject to further testing and refinement" (emphasis in original)). But, in order to qualify as "scientific knowledge," an inference or assertion must be derived by the scientific method. Proposed testimony must be supported by appropriate validation—*i.e.*, "good grounds," based on what is known. In short, the requirement that an expert's testimony pertain to "scientific knowledge" establishes a standard of evidentiary reliability. . . .

Faced with a proffer of expert scientific testimony, then, the trial judge must determine at the outset, pursuant to Rule 104(a),[10] whether the expert is proposing to testify to (1) scientific knowledge that (2) will assist the trier of fact to understand or determine a fact in issue.[11] This entails a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid and of whether that reasoning or methodology properly can be applied to the facts in issue. We are confident that federal judges possess the capacity to undertake this review. Many factors will bear on the inquiry, and we do not presume to set out a definitive checklist or test. But some general observations are appropriate.

Ordinarily, a key question to be answered in determining whether a theory or technique is scientific knowledge that will assist the trier of fact will be whether it can be (and has been) tested. "Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified; indeed, this methodology is what distinguishes science from other fields of human inquiry." Green, Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of *Agent Orange* and Bendectin Litigation, 86 Nw. U. L. Rev. 643, 645 (1992). See also C. Hempel, Philosophy of Natural Science 49 (1966) ("The statements constituting a scientific explanation must be capable of empirical test"); K. Popper, Conjectures and Refutations: The Growth of Scientific Knowledge 37 (5th ed. 1989) ("The criterion of the scientific status of a theory is its falsifiability, or refutability, or testability") (emphasis deleted).

10. Rule 104(a) provides:

> "Preliminary questions concerning the qualification of a person to be a witness, the existence of a privilege, or the admissibility of evidence shall be determined by the court, subject to the provisions of subdivision (b) [pertaining to conditional admissions]. In making its determination it is not bound by the rules of evidence except those with respect to privileges."

These matters should be established by a preponderance of proof.

11. Although the *Frye* decision itself focused exclusively on "novel" scientific techniques, we do not read the requirements of Rule 702 to apply specially or exclusively to unconventional evidence. Of course, well-established propositions are less likely to be challenged than those that are novel, and they are more handily defended. Indeed, theories that are so firmly established as to have attained the status of scientific law, such as the laws of thermodynamics, properly are subject to judicial notice under Federal Rule of Evidence 201.

Another pertinent consideration is whether the theory or technique has been subjected to peer review and publication. Publication (which is but one element of peer review) is not a *sine qua non* of admissibility; it does not necessarily correlate with reliability, see S. Jasanoff, The Fifth Branch: Science Advisors as Policymakers 61–76 (1990), and in some instances well-grounded but innovative theories will not have been published, see Horrobin, The Philosophical Basis of Peer Review and the Suppression of Innovation, 263 JAMA 1438 (1990). Some propositions, moreover, are too particular, too new, or of too limited interest to be published. But submission to the scrutiny of the scientific community is a component of "good science," in part because it increases the likelihood that substantive flaws in methodology will be detected. See J. Ziman, Reliable Knowledge: An Exploration of the Grounds for Belief in Science 130–133 (1978); Relman & Angell, How Good Is Peer Review?, 321 New Eng. J. Med. 827 (1989). The fact of publication (or lack thereof) in a peer reviewed journal thus will be a relevant, though not dispositive, consideration in assessing the scientific validity of a particular technique or methodology on which an opinion is premised.

Additionally, in the case of a particular scientific technique, the court ordinarily should consider the known or potential rate of error, see, e.g., *United States* v. *Smith*, 869 F.2d 348, 353–354 (CA7 1989) (surveying studies of the error rate of spectrographic voice identification technique), and the existence and maintenance of standards controlling the technique's operation, see *United States* v. *Williams*, 583 F.2d 1194, 1198 (CA2 1978) (noting professional organization's standard governing spectrographic analysis), cert. denied, 439 U.S. 1117, 59 L. Ed. 2d 77, 99 S. Ct. 1025 (1979).

Finally, "general acceptance" can yet have a bearing on the inquiry. A "reliability assessment does not require, although it does permit, explicit identification of a relevant scientific community and an express determination of a particular degree of acceptance within that community." *United States* v. *Downing*, 753 F.2d at 1238. . . . Widespread acceptance can be an important factor in ruling particular evidence admissible, and "a known technique which has been able to attract only minimal support within the community," *Downing*, 753 F.2d at 1238, may properly be viewed with skepticism. . . .

Chief Justice Rehnquist, with whom Justice Stevens joins, concurring in part and dissenting in part.

[The dissenters concurred in the judgment that the *Frye* test should be superseded, but disagreed with some of the majority's "general observations" about how the new standard should be applied.] The Court speaks of its confidence that federal judges can make a "preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid and of whether that reasoning or methodology properly can be applied to the facts in issue." *Ante*, at 592-593. The Court then states that a "key question" to be answered in deciding whether something is "scientific knowledge" "will be whether it can be (and has been) tested." *Ante*, at 593. Following this sentence are three quotations from treatises, which not only speak of empirical testing, but one of which states that the " 'criterion of the scientific status of a theory is its falsifiability, or refutability, or testability.' " *Ibid.*

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.

I do not doubt that Rule 702 confides to the judge some gatekeeping responsibility in deciding questions of the admissibility of proffered expert testimony. But I do not think it imposes on them either the obligation or the authority to become amateur scientists in order to perform that role. I think the Court would be far better advised in this case to decide only the questions presented, and to leave the further development of this important area of the law to future cases.

NOTES AND QUESTIONS

1. *Daubert* holds that, to satisfy Rule 702, "[t]he subject of an expert's testimony must be 'scientific . . . knowledge,' " and that "in order to qualify as 'scientific knowledge,' an inference or assertion must be derived by the scientific method." Thus, Rule 702 requires trial judges to determine whether the inferences and assertions of scientific experts are derived from the application of the scientific method. But how are they to know the scientific method when they see it? In other words, what is the Supreme Court's philosophy of science?

2. Although the Court did "not presume to set out a definitive checklist or test," it did presume, under the guise of "general observations," to set out what have become famous as the "four *Daubert* factors": (1) whether the theory or technique "can be (and has been) tested," or falsified; (2) "whether the theory or technique has been subjected to peer review and publication"; (3) "the known or potential rate of error" of the technique; and (4) general acceptance, demoted from controlling factor to secondary consideration. One widespread practical criticism of these factors is that they call on trial judges to do too much. Whereas the majority was "confident that federal judges possess the capacity to undertake this review," the dissent was skeptical about trial judges functioning as "amateur scientists." To illustrate the problem, Chief Justice Rehnquist observed, "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." (If he had thought about it further, he might have been even more confused about what "error rate" means—see Section II.B *infra*.) There continues to be controversy over who was right. As frequent teachers of judges, the authors can affirm that the trial bench takes the *Daubert* duty seriously and works hard at it. Practicing

lawyers and testifying experts report wide variation in judges' ability to know the scientific method when they see it. Early in the *Daubert* era, one federal district judge observed, somewhat sourly, that "[t]hings have changed in recent years. The language problem has expanded and not merely because world commerce brings more languages into our courts. Rather, there are dozens of areas of scientific and technical expertise, and those who offer such testimony often speak in the functional equivalent of Urdu, and translation is impossible without understanding some principles of the relevant science." United States ex rel. Kokoraleis v. Director of the Ill. Dep't of Corrections, 963 F. Supp. 1473, 1488 (N.D. Ill. 1997) (Zagel, J.).

3. A second, more substantive criticism of the *Daubert* factors is that they are too narrow. The contention is that the four factors properly apply only to science on the physics model, in which (at least ideally; see the next reading) hypotheses are developed and then tested in rigorous experiments—what is often called positivist, or "hard" science. Consequently, it is argued, the strict application of the factors may lead to the rejection of many kinds of expert analysis that fail to meet the standards of physics but are nonetheless "scientific" in the commonly understood meaning of the term (for example, the testimony of clinical physicians about the causes of disease, discussed in Chapter 7, Medical Causation). The rest of this chapter is devoted to the exploration of this critique. We will explore alternative conceptions of the meaning of science, beginning with those of scientists themselves and then turning to those expressed by judges. The goal is to develop a general understanding of science and the scientific method that can serve as background for the study of specific scientific disciplines in the subsequent chapters.

II. HOW SCIENTISTS THINK ABOUT SCIENCE

A. The Philosophy of Science

How Science Works

Reference Manual on Scientific Evidence 39–45 (Federal Judicial Center, 3d ed. 2011) David Goodstein

III. Theories of Science. . .

A. francis bacon's scientific method

. . . Ask a scientist what science is, and the answer will almost surely be that it is a process—a way of examining the natural world and discovering important truths about it. In short, the essence of science is the scientific method.

That stirring description suffers from an important shortcoming. We don't really know what the scientific method is. There have been many attempts at formulating a general theory of how science works, or at least how it should work, starting, as we have seen, with the theory of Sir Francis Bacon. But Bacon's idea, that science proceeds through the collection of observations without prejudice, has been rejected by all serious thinkers. Everything about the way we do science — the language we use, the instruments we use, the methods we use — depends on clear presuppositions about how the world works. Modern science is full of things that cannot be observed at all, such as force fields and complex molecules. At the most fundamental level, it is impossible to observe nature without having some reason to choose what is worth observing and what is not worth observing. Once that elementary choice is made, Bacon has been left behind.

b. Karl Popper's Falsification Theory

Over the past century, the ideas of the Vienna-born philosopher Sir Karl Popper have had a profound effect on theories of the scientific method. In contrast to Bacon, Popper believed that all science begins with a prejudice, or perhaps more politely, a theory or hypothesis. Nobody can say where the theory comes from. Formulating the theory is the creative part of science, and it cannot be analyzed within the realm of philosophy. However, once the theory is in hand, Popper tells us, it is the duty of the scientist to extract from it logical but unexpected predictions that, if they are shown by experiment not to be correct, will serve to render the theory invalid.

Popper was deeply influenced by the fact that a theory can never be proved right by agreement with observation, but it can be proved wrong by disagreement with observation. Because of this asymmetry, science uniquely makes progress by proving that good ideas are wrong so that they can be replaced by even better ideas. Thus, Bacon's impartial observer of nature is replaced by Popper's skeptical theorist. The good Popperian scientist somehow comes up with a hypothesis that fits all or most of the known facts, then proceeds to attack that hypothesis at its weakest point by extracting from it predictions that can be shown to be false. This process is known as falsification.

Popper's ideas have been fruitful in weaning the philosophy of science away from the Baconian view and some other earlier theories, but they fall short in a number of ways in describing correctly how science works. The first of these is the observation that, although it may be impossible to prove a theory is true by observation or experiment, it is as almost equally impossible to prove one is false by these same methods. Almost without exception, in order to extract a falsifiable prediction from a theory, it is necessary to make additional assumptions beyond the theory itself. Then, when the prediction turns out to be false, it may well be one of the other assumptions, rather than the theory itself, that is false. To take a simple example, early in the twentieth century it was found that the orbits of the outermost planets did not quite obey the predictions of Newton's laws of gravity and mechanics. Rather than take this to be a falsification of Newton's laws, astronomers concluded that the orbits were being perturbed by an additional unseen body out there. They were right. That is precisely how Pluto was discovered.

The apparent asymmetry between falsification and verification that lies at the heart of Popper's theory thus vanishes. But the difficulties with Popper's view go even beyond that problem. It takes a great deal of hard work to come up with a new theory that is consistent with nearly everything that is known in any area of science. Popper's notion that the scientist's duty is then to attack that theory at its most vulnerable point is fundamentally inconsistent with human nature. It would be impossible to invest the enormous amount of time and energy necessary to develop a new theory in any part of modern science if the primary purpose of all that work was to show that the theory was wrong.

This point is underlined by the fact that the behavior of the scientific community is not consistent with Popper's notion of how it should be. Credit in science is most often given for offering correct theories, not wrong ones, or for demonstrating the correctness of unexpected predictions, not for falsifying them. I know of no example of a Nobel Prize awarded to a scientist for falsifying his or her own theory.

C. Thomas Kuhn's Paradigm Shifts

Another towering figure in the twentieth century theory of science is Thomas Kuhn. Kuhn was not a philosopher but a historian (more accurately, a physicist who retrained himself as a historian). It is Kuhn who popularized the word *paradigm*, which has today come to seem so inescapable.

A paradigm, for Kuhn, is a sort of consensual worldview within which scientists work. It comprises an agreed-upon set of assumptions, methods, language, and everything else needed to do science. Within a given paradigm, scientists make steady, incremental progress, doing what Kuhn calls "normal science."

As time goes on, difficulties and contradictions arise that cannot be resolved, but the tendency among scientists is to resist acknowledging them. One way or another, they are swept under the rug, rather than being allowed to threaten the central paradigm. However, at a certain point, enough of these difficulties accumulate to make the situation intolerable. At that point, a scientific revolution occurs, shattering the paradigm and replacing it with an entirely new one.

This new paradigm, says Kuhn, is so radically different from the old that normal discourse between the practitioners of the two paradigms becomes impossible. They view the world in different ways and speak different languages. It is not even possible to tell which of the two paradigms is superior, because they address different sets of problems. They are incommensurate. Thus, science does not progress incrementally, as the science textbooks would have it, except during periods of normal science. Every once in a while, a scientific revolution brings about a paradigm shift, and science heads off in an entirely new direction.

Kuhn's view was formed largely on the basis of two important historical revolutions. One was the original scientific revolution that started with Nicolaus Copernicus and culminated with the new mechanics of Isaac Newton. The very word *revolution*, whether it refers to the scientific kind, the political kind, or any other kind, refers metaphorically to the revolutions in the heavens that Copernicus described in a book, *De Revolutionibus Orbium Caelestium*, which was published as he lay dying in 1543. Before Copernicus, the dominant paradigm was the world view of ancient Greek philosophy, frozen in the fourth century B.C.E. ideas of Plato and Aristotle. After Newton, whose masterwork, *Philosophiae Naturalis Principia Mathematica*, was published in 1687, every scientist was a Newtonian, and Aristotelianism was banished forever from the world stage. It is even possible that Sir Francis Bacon's disinterested observer was a reaction to Aristotelian authority. Look to nature, not to the ancient texts, Bacon may have been saying.

The second revolution that served as an example for Kuhn occurred early in the twentieth century. In a headlong series of events that lasted a mere 25 years, the Newtonian paradigm was overturned and replaced with the new physics, in the form of quantum mechanics and Einstein's theories of special and general relativity. This second revolution, although it happened much faster, was no less profound than the first.

The idea that science proceeds by periods of normal activity punctuated by shattering breakthroughs that make scientists rethink the whole problem is an appealing one, especially to the scientists themselves, who know from personal experience that it really happens that way. Kuhn's contribution is important. It offers us a useful context (a paradigm, one might say) for organizing the entire history of science.

Nonetheless, Kuhn's theory does suffer from a number of shortcomings. One of them is that it contains no measure of how big the change must be in order to qualify as a revolution or paradigm shift. Most scientists will say that there is a paradigm shift in their laboratory every 6 months or so (or at least every time it becomes necessary to write another proposal for research support). That is not exactly what Kuhn had in mind.

Another difficulty is that even when a paradigm shift is truly profound, the paradigms it separates are not necessarily incommensurate. The new sciences of quantum mechanics and relativity, for example, did indeed show that Newton's laws of mechanics were not the most fundamental laws of nature. However, they did not show that they were wrong. Quite the contrary, they showed why Newton's laws were right: Newton's laws arose out of newly discovered laws that were even deeper and that covered a wider range of circumstances unimagined by Newton and his followers—that is, phenomena as small as atoms, or nearly as fast as the speed of light, or as dense as black holes. In our more familiar realms of experience, Newton's laws go on working just as well as they always did. Thus, there is no quarrel and no ambiguity at all about which paradigm is "better." The new laws of quantum mechanics and relativity subsume and enhance the older Newtonian worldview.

D. AN EVOLVED THEORY OF SCIENCE

If neither Bacon nor Popper nor Kuhn gives us a perfect description of what science is or how it works, all three of them help us to gain a much deeper understanding of it.

Scientists are not Baconian observers of nature, but all scientists become Baconians when it comes to describing their observations. With very few exceptions, scientists are rigorously, even passionately, honest about reporting scientific results and how they were obtained. Scientific data are the coin of the realm in science, and they are always treated with reverence. Those rare instances in which scientists are found to have fabricated or altered their data in some way are always traumatic scandals of the first order.

Scientists are also not Popperian falsifiers of their own theories, but they do not have to be. They do not work in isolation. If a scientist has a rival with a different theory of the same phenomena, the rival will be more than happy to perform the Popperian duty of attacking the scientist's theory at its weakest point. Moreover, if falsification is no more definitive than verification, and scientists prefer in any case to be right rather than wrong, they nonetheless know how to hold verification to a very high standard. If a theory makes novel and unexpected predictions, and those predictions are verified by experiments that reveal new and useful or interesting phenomena, then the chances that the theory is correct are greatly enhanced. And, even if it is not correct, it has been fruitful in the sense that it has led to the discovery of previously unknown phenomena that might prove useful in themselves and that will have to be explained by the next theory that comes along.

Finally, science does not, as Kuhn seemed to think, periodically self-destruct and need to start over again. It does, however, undergo startling changes of perspective that lead to new and, invariably, better ways of understanding the world. Thus, although science does not proceed smoothly and incrementally, it is one of the few areas of human endeavor that is genuinely progressive. There is no doubt at all that twentieth century science is better than nineteenth century science, and we can be absolutely confident that the quality of science in the twenty-first century will be better still. One cannot say the same about, say, art or literature.

To all of this, a few things must be added. The first is that science is, above all, an adversarial process. It is an arena in which ideas do battle, with observations and data the tools of combat. The scientific debate is very different from what happens in a court of law, but just as in the law, it is crucial that every idea receive the most vigorous possible advocacy, just in case it might be right. Thus, the Popperian ideal of holding one's hypothesis in a skeptical and tentative way is not merely inconsistent with reality; it would be harmful to science if it were pursued. As will be discussed shortly, not only ideas, but the scientists themselves, engage in endless competition according to rules that, although they are not written down, are nevertheless complex and binding.

In the competition among ideas, the institution of peer review plays a central role. Scientific articles submitted for publication and proposals for funding are often sent to anonymous experts in the field, in other words, peers of the author, for review. Peer review works superbly to separate valid science from nonsense, or, in Kuhnian terms, to ensure that the current paradigm has been respected.[11] It works less well as a means of choosing between competing valid ideas, in part because the peer doing the reviewing is often a competitor for the same resources (pages in prestigious journals, funds from government agencies or private foundations) being sought by the authors. It works very poorly in catching cheating or fraud, because all scientists are socialized to believe that even their toughest competitor is rigorously honest in the reporting of scientific results, which makes it easy for a purposefully dishonest scientist to fool a referee. Despite all of this, peer review is one of the venerated pillars of the scientific edifice.

NOTES AND QUESTIONS

1. Goodstein is a physics professor and senior administrator at Caltech, and thus at the center of "hard" scientific research. His article is a remarkably concise survey of the history and philosophy of science. Which of the

11. The Supreme Court received differing views regarding the proper role of peer review. *Compare* Brief for Amici Curiae Daryl E. Chubin et al. at 10, Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579 (1993) (No. 92-102) ("peer review referees and editors limit their assessment of submitted articles to such matters as style, plausibility, and defensibility; they do not duplicate experiments from scratch or plow through reams of computer-generated data in order to guarantee accuracy or veracity or certainty"), *with* Brief for Amici Curiae New England Journal of Medicine, Journal of the American Medical Association, and Annals of Internal Medicine in Support of Respondent, Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579 (1993) (No. 92-102) (proposing that publication in a peer-reviewed journal be the primary criterion for admitting scientific evidence in the courtroom). *See generally* Daryl E. Chubin & Edward J. Hackett, Peerless Science: Peer Review and U.S. Science Policy (1990); Arnold S. Relman & Marcia Angell, *How Good Is Peer Review?* 321 New Eng. J. Med. 827–29 (1989). As a practicing scientist and frequent peer reviewer, I can testify that Chubin's view is correct.

theories of science that Goodstein describes did the Supreme Court seem to adopt in *Daubert*? Here is Goodstein's answer to that question:

> In reading these four illustrative criteria mentioned by the Court, one is struck immediately by the specter of Karl Popper looming above the robed justices. (It's no mere illusion. The dependence on Popper is explicit in the written decision.) Popper alone is not enough, however, and the doctrine of falsification is supplemented by a bow to the institution of peer review, an acknowledgment of the scientific meaning of error, and a paradigm check (really, an inclusion of the earlier *Frye* standard).
>
> All in all, I would give the decision pretty high marks. The justices ventured into the treacherous crosscurrents of the philosophy of science—where even most scientists fear to tread—and emerged with at least their dignity intact. Falsifiability may not be a good way of doing science, but it's not the worst a posteriori way to judge science, and that's all that's required here. At least they managed to avoid the Popperian trap of demanding that the scientists be skeptical of their own ideas. The other considerations help lend substance and flexibility. The jury is still out (so to speak) on how well this decision will work in practice, but it is certainly an impressive attempt to serve justice, if not truth.

Id. at 53–54.

2. The word "positivism" is often associated with the Popperian theory of science. Goodstein suggests that it is generally inadequate to account for how science is actually practiced. But is it a better fit for some scientific enterprises than others? In considering the adequacy of the positivist model, would it make sense to distinguish Goodstein's own physics from, say, clinical psychology? Is it a question of degree—are some sciences more Popperian than others? We return to this question in Section III *infra*.

3. Other scientists are more sanguine about achieving the Popperian ideal than Goodstein. They tend to see a greater divide between scientific and legal approaches to proof, and to take a less generous view of the latter. See, for example, the New England Journal of Medicine's post-*Daubert* editorial, which discusses the controversial litigation over the safety of breast implants:

> [W]hile the FDA was considering whether to remove breast implants from the market, public concern was mounting. When the implants were banned, the concern did not abate, because of the many women who already had breast implants. Frightening stories about the effects of breast implants swept through the media and were reified by repetition. The fact that implants occasionally leaked or ruptured, causing local problems, was well known, but the new stories focused primarily on anecdotes of serious systemic effects. Breast implants were said to cause arthritis or polymyositis or systemic lupus erythematosus. Since these are chronic disorders that are not always

easily diagnosed and since they clearly can occur in women who do not have breast implants as well as in those who do, the postulated associations were very difficult to evaluate, and until now there has been little attempt to do so systematically.

Despite the lack of published epidemiologic studies, the accumulated weight of anecdotes was taken by judges and juries as tantamount to proof of causation. Multimillion-dollar settlements followed, along with poignant stories in the media and appearances by plaintiffs on talk shows. All this added to the weight of the anecdotes, which in a circular way became accepted by the courts and the public as nearly incontrovertible evidence. Three manufacturers of breast implants finally decided that a lump settlement would be less expensive than to go on losing cases one by one, so they agreed to a class-action settlement, establishing a fund of $4.2 billion. . . .

What does this tell us about the way scientific issues are settled in the courtroom? As readers of the Journal know, scientific conclusions cannot be based on argument and opinion. There must be data. Yet, in the courtroom, acceptance of expert testimony on scientific questions usually turns on the "credibility" of the witness, not the validity of the evidence on which the witness's opinion is based. Furthermore, expert witnesses are selected by the contesting lawyers, often paid by them, and their testimony is reviewed in advance — circumstances unlikely to ensure objectivity or even competence. The resulting judgments are sometimes manifestly ludicrous. For example, one court awarded damages to a woman who claimed that she lost her psychic powers because of a computed tomographic scan. (This decision was later overturned.)

Marcia Angell, M.D., Editorial: *Do Breast Implants Cause Systemic Disease?— Science in the Courtroom*, 330 New England J. Med. 1748 (1994).*

B. The Scientific Ideal in Practice

Of Cherries, Fudge, and Onions: Science and Its Courtroom Perversion

64 Law & Contemporary Problems 213 (2001)
David W. Peterson and John M. Conley

II. A Scientific Ideal: The Designed Experiment

A. THE SALK VACCINE TRIALS

As of the early 1950s, the poliomyelitis virus was a scourge of America's children; hundreds of thousands were afflicted, and many were disabled for life. Among several vaccines under development, the one produced

* Reprinted with permission of Massachusetts Medical Society and the author. — Eds.

by Dr. Jonas Salk showed considerable promise. The United States Public Health Service decided to conduct a very large-scale experiment to determine its effectiveness. Ultimately, about two million school children were involved in the tests, though only a fraction of them actually received the vaccine.

In designing the experiment, the Public Health Service attempted to take account of a variety of special considerations. The first was the nature of the disease itself. Polio is a hygiene-related disease: children who live in unhygienic environments are more likely to contract polio than those who live in cleaner environments. Perversely, it is the latter children who are most severely affected when they do contract the disease; the former are likely to be exposed early to the polio virus, so they tend to suffer only mildly and to develop an immunity to further harm. Polio is also contagious, so that if one second-grader is infected, there is an increased chance that his or her classmates will be infected. Furthermore, polio is epidemic, so the incidence is much greater in some years than in others.

Second, there were ethical and technical considerations distinct from the nature of the disease. Clearly, one could not ethically require that any particular child be vaccinated with the experimental vaccine without permission from the child's parents or guardian. But it is possible that children whose parents would grant permission would differ in some material and systematic ways from children whose parents would withhold permission. For example, it might be that well-educated and relatively affluent parents would tend to grant permission, while less well-educated and less affluent parents would not. A potential result is that relatively many children living in hygienic circumstances would be permitted to take part in the study, and relatively few of those living in less hygienic conditions. This imbalance could seriously skew the results of the study.

There is also the problem that the behavior of the child or the parents might be influenced by the fact that the child had been vaccinated. A child thus protected need not be quite as cautious in avoiding possible exposure to polio, and therefore might tend to engage in riskier behavior than his non-vaccinated neighbor. This too could seriously distort the study results. Moreover, since polio comes in both mild and severe forms, it is not always clear whether a child has contracted the virus. As a result, a clinician examining a child who had been vaccinated might be less inclined to diagnose polio for that child than for her unvaccinated neighbor.

Sorting through this web of considerations, the Public Health Service decided upon the following course of action. First, it selected schools across the nation where the incidence of polio was relatively high. It then sought the permission of parents of first, second, and third graders for their children to be vaccinated as part of the study. Half of the participating children were selected *at random* and injected with the Salk vaccine. The other half of the subjects were injected with a placebo, a saline solution designed to have no medical effect whatsoever. The children and guardians did not know whether a child received the vaccine or the placebo. All of the children—both those who received the Salk vaccine and those who received the placebo—were monitored over the ensuing months by clinicians who also were not told which of the children had received which treatment.

An experiment of this sort is termed randomized and double-blind. It is randomized because the choice of whom to give the real vaccine is made by the toss of a coin or some other equally detached chance process. This process virtually guarantees that there will be no systematic difference between the group of children given the real vaccine and those who are given the false vaccine. It is blind in the first instance because the children and their parents are unaware of whether they have actually been vaccinated. Consequently, there is virtually no chance that the vaccinated children, as a group, will behave any differently from those who were given the false vaccine. It is blind in the second instance because the people evaluating the health of the children do not know which of their subjects received the Salk vaccine and which received the placebo. As a result, it is virtually certain that the same methods and standards for diagnosis will be used for the vaccinated group as for the placebo group.

In all, about 200,000 students received the Salk vaccine and about 200,000 received the placebo in this phase of the experiment. The incidence of polio among the vaccinated group was approximately twenty-eight cases per 100,000, while that among the placebo group was about seventy-one cases per 100,000. Given the randomized, double-blind construction of this experiment, there are only two possible explanations for these results. The first is that the Salk vaccine really differed from the placebo in its effect on polio and that the difference was in the direction of reducing polio. The second is that the Salk vaccine was no different from the placebo in its effect on polio, and that the observed reduction in the incidence in polio was due solely to the manner in which children were assigned to the treatment groups. In other words, the children who contracted polio were destined to get it regardless of their treatment and the fact that most of them were placed in the group given the placebo was due purely to the luck of the coin toss. . . .

Part of the genius of this experimental design is that one can calculate the probability that, if indeed the Salk vaccine were identical to the placebo in its effect on the polio virus, the coin toss mechanism would result in so many of the children predestined to contract polio being assigned to the placebo group. That probability is about one in a billion. Thus, it seems safe to say that the difference in the incidence of polio between the Salk group and the placebo group cannot reasonably be attributed to the random assignment of children to treatment and placebo groups. The only remaining possibility is that the difference was caused by the greater effectiveness of the Salk vaccine. . . .

This is the logic of the ideal scientific experiment. The design features of random assignment and double-blinding virtually rule out the possibility of systematic differences between the experimental and control groups other than exposure to the suspected causal agent. When [statistical testing] indicates that chance is too unlikely an explanation for an observed disparity in outcomes, the only alternative is to conclude that the agent's causal effect is real.

The problem is that few questions, scientific or otherwise, can be settled with the elegant finality of the Salk trials. This is particularly true in legal contexts, where the evidence is almost never so neat. In many instances, logistical or ethical barriers preclude a true experiment. Researchers are relegated to uncontrolled observational studies or after-the-fact data analysis. In all such cases, the focus on the suspected causal agent can never be as sharp as in the ideal, well-designed experiment.

These situations have proved to be particularly troubling for courts under the *Daubert* regime. When the only scientific evidence of causation falls short of the Salk trials' gold standard, as it does in most cases, does it comport sufficiently with the scientific method to be admissible? That is, is it sufficiently reliable to be translated into legal causation? This is precisely the question that troubled the courts that heard the *Daubert* case itself as they wrestled with the issue of whether Bendectin could be reliably shown to cause birth defects.

NOTES AND QUESTIONS

1. Do you see why the authors describe the Salk vaccine trials as the scientific "gold standard"? Does this research at least approximate the Popperian ideal? What was the researchers' hypothesis? How did they test it? Why did they reach the conclusion that they did? As you will see in Chapter 3, the methods of statistical inference introduce some complexities into the hypothesis-testing process that are not discussed in this excerpt.

2. Assume that a qualified scientist wished to testify in a federal court that the Salk vaccine is effective in preventing polio. Assume that the results of the vaccine trials had been published in a peer-reviewed journal. How would the expert's assertion fare under the other three *Daubert* factors? Has it been tested/falsified? What might "error rate" mean here? How would you assess "general acceptance"?

3. As Peterson and Conley contend, the "gold standard" is rarely achieved in science, and particularly the science that experts regularly testify about. As the next section illustrates, scientific evidence offered in court is often less Popperian. As the next two cases illustrate, judges have recognized this problem and struggled to deal with it.

III. HOW THE COURTS UNDERSTAND SCIENCE: IRRECONCILABLE DIFFERENCES?

The problem of dealing with scientific evidence is not new:

> The federal judiciary's love-hate relationship with scientific evidence dates back to at least 1908. In that year, in *Muller v. Oregon*,[1] the Supreme Court received the eponymous "Brandeis brief." Louis Brandeis, defending an Oregon statute that limited the working hours of women, submitted "a very copious collection" of authorities purporting to show the particular vulnerability of women in the turn-of-the-century workplace. The Court upheld the law, commenting on Brandeis' submission with sibylline brevity: "It may not be amiss, in the present case, before examining the constitutional question, to notice the course of legislation as well as expressions of opinion from other than judicial sources" Drawing on Brandeis' brief, the Court found "[t]hat woman's physical structure and the performance of maternal functions place her at a disadvantage in the struggle for subsistence."
>
> Since *Muller*, the courts' attitudes toward science in the courtroom have run the gamut from uncritical enthusiasm to dismissive Luddism, with stops at all intermediate points. Just nineteen years after *Muller*, Oliver Wendell Holmes, Jr. led the Supreme Court into the depths of evangelical credulity with his infamous opinion in *Buck v. Bell* upholding the compulsory sterilization of "imbeciles" on the basis of turn-of-the-century theories of intelligence testing.[5] A generation later, in footnote eleven of its opinion in *Brown v. Board of Education*, the Court turned to science to support its finding of constitutional fact that separate education is inherently unequal.[6] In the 1970s, as employment discrimination litigation proliferated, the Court plunged into detailed questions of scientific method, endorsing particular tests of the statistical significance of racial disparities in hiring.[7] Over the last ten years, particularly in its death penalty jurisprudence, the Court has displayed a more skeptical attitude toward scientific evidence. Sometimes this skepticism has expressed itself in the form of detailed critiques of particular research;[8] at other times, Court majorities have questioned whether broad-based scientific studies can ever be probative

1. 208 U.S. 412 (1908).
5. 274 U.S. 200, 207 (1927).
6. 347 U.S. 483, 494 n.11 (1954).
7. Hazelwood Sch. Dist. v. United States, 433 U.S. 299, 308 n.14 (1977) (endorsing the proposition that employer's denial of discrimination will be "suspect" where representation of protected group in employer's workforce is more than two or three standard deviations below what would be expected on basis of population data); Castenada v. Partida, 430 U.S. 482, 466 n.17 (1976) (same).
8. *See, e.g.*, Lockhart v. McCree, 476 U.S. 162, 170–73 (1976) (criticizing studies intended to show bias of guilt-phase juries in capital cases).

of individual constitutional violations.[9] Categorizing trends in the lower court's reception of scientific evidence would be a book-length undertaking. For background purposes, suffice it to say that one can find case support for almost any side of nearly every scientific question that has ever come before the courts.

John M. Conley and David W. Peterson, *The Science of Gatekeeping: The Federal Judicial Center's New Reference Manual on Scientific Evidence*, 74 N.C.L. Rev. 1183, 1184-1185 (1996).

The first of the two cases that follow lays out a strongly Popperian understanding of "hard" science, and then attempts to categorize the work of the practicing physician. The second discusses the special problems of dealing with the "soft," statistically based social sciences. Be sure to consult the notes after the cases, as each has an unusual and significant history, and each is treated in further detail in subsequent chapters.

Moore v. Ashland Chemical, Inc.

126 F.3d 679 (5th Cir. 1997), *superseded by* 151 F.3d 269 (5th Cir. 1998) (en banc)

[Moore claimed that his rare lung disease was caused by exposure to toluene on the job and sued his employer. In attempting to prove causation, he relied primarily on the opinions of the clinical physicians who cared for him. The district court rejected these opinions under *Daubert*. In this, its initial opinion, the Fifth Circuit reversed, holding that the district court had abused its discretion in excluding the doctors' testimony.]

[C]linical medicine (as opposed to research and laboratory medical science) is not a hard science discipline; its goals, subject matter, conditions of study and well developed methodology are sui generis and quite different from that of hard science and its methodology. . . .

A. Hard Scientific Knowledge

Speaking specifically of "scientific knowledge," the [*Daubert*] Court stated that the adjective " 'scientific' implies a grounding in the methods and procedures of science." 509 U.S. at 592. The Court elaborated:

> "Science is not an encyclopedic body of knowledge about the universe. Instead, it represents a *process* for proposing and refining theoretical explanations about the world that are subject to further testing and refinement. . . ." But, in order to qualify as "scientific knowledge," an inference or assertion must be derived by the scientific method. . . .

9. *See, e.g.*, McClesky v. Kemp, 481 U.S. 279, 292–95 (1987) (questioning whether social science data can ever prove sentencing to be racially discriminatory in any particular case).

Thus, the *Daubert* Court defined "scientific knowledge" in terms of "hard science" or "Newtonian science" i.e., knowledge obtained and tested through "the scientific method," of which Sir Isaac Newton was the leading exponent. . . .

The methodology of hard or Newtonian science is what distinguishes it from other fields of human inquiry. See Michael D. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of Agent Orange and Bendectin Litigation*, 86 Nw. U.L. Rev. 643, 645 (1992). "Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified. . . . Theoretically, therefore, hypotheses are not affirmatively proved, only falsified. Of course, if a hypothesis repeatedly withstands falsification, one may tend to accept it even if conditionally true." *Id.* at 645-646 (citing Karl R. Popper, *The Logic of Scientific Discovery* (1965); David L. Faigman, *To Have and Have Not: Assessing the Value of Social Science to the Law as Science and Policy*, 38 Emory L.J. 1005, 101517 (1989). . . .

In *Daubert*, the Court indicated that, (1) "scientific knowledge" within Rule 702 means principles, theories, techniques, or inferences derived by the scientific method or by a body of sound scientific methods; and (2) that the proffered expert's opinion, inference, or testimony based on scientific knowledge, in order to have evidentiary reliability or trustworthiness, must be derived or inferred by the same methods. *Id.* at 590 n.9; *see also* the court's "general observations" on principal scientific methods. *Id.* at 593–594.

By the same token, we conclude that, under Rule 702, an opinion based on other technical or specialized knowledge, must be grounded in the principles, methods, and procedures of the particular field of knowledge involved. Every discipline employs a body of methods, rules, and postulates, i.e., methodology, both in its ordinary functions and in developing and adopting new concepts, techniques, and analogues. Therefore, the "knowledge" of each discipline, under Rule 702, is both its principles and methodology and the theories, techniques, or inferences produced through its methodology. Thus, the proffered opinion of any expert in a field of knowledge, in order to be evidentiarily reliable, must either be based soundly on the current knowledge, principles, and methodology of the expert's discipline or be soundly inferred or derived therefrom. . . .

F. The *Daubert* "Factors" Are Hard Scientific Methods That Generally Are Inappropriate for the Reliability Assessment of Clinical Medical Testimony.

After declaring that evidentiary reliability of an expert's scientific opinion depends on whether it is soundly grounded in the scientific method, the *Daubert* Court identified several individual methods or techniques within the body of hard or Newtonian scientific methodology as appropriate for trial judges' use in testing the methodology-relatedness of particular hard scientific opinion proffers. *Daubert,* 509 U.S. at 593. These hard scientific methods, now sometimes called "*Daubert* factors," are empirical testing, peer review and publication, known or potential rate of error, the existence and maintenance of operational standards, and acceptance within a relevant scientific community. *Id.* at 593–594.

Because the objectives, functions, subject matter, and methodology of hard science vary significantly from those of the discipline of clinical medicine, as distinguished from research or laboratory medicine, the hard science techniques or methods that became the "*Daubert* factors" generally are not appropriate for assessing the evidentiary reliability of a proffer of expert clinical medical testimony.

First, the goals of the disciplines of clinical medicine and hard or Newtonian science are different. In hard science, the usual motive is inquiring: to gain a new understanding of some mechanism of nature. Alvan R. Feinstein, *Clinical Judgment* 22 (1967) [hereinafter Feinstein]. In contrast, the care and treatment of the individual patient is the ultimate, specific act that characterizes a clinical physician. *Id.* at 27; Pellegrino and Thomasma, *For The Patient's Good* 71 (1988); Pellegrino and Thomasma, *A Philosophical Basis of Medical Practice* 120 (1981) ("The whole process is ordained to a specific practical end—a right action for a particular patient—and . . . this end must modulate each step leading to it in important ways."). The clinical physician, therefore, must take account of the immediacy of the problem confronting her for she bears an essential relationship to each patient. Additionally, she has many human values to consider—ethics, compassion, and must have a willingness to take responsibility in the face of the unknown. Edmond A. Murphy, *The Logic of Medicine* 6 (1976) [hereinafter Murphy]. The pursuit of these different goals—of hard science and clinical medicine—serves to shape the distinct objectives of the scientific experiment and the clinical treatment of a patient:

> In clinical treatment, the main motives are remedial, or prophylactic: to change what nature has done or to prevent what it may do. In laboratory work, the premise is innovative: the goal is to test a new hypothesis or a new procedure. In ordinary clinical treatment, the premise is repetitive: the goal is to reproduce (or surpass) the best results of experiments conducted before in similar circumstances. A clinician chooses treatment in a new situation by reviewing what was done and what happened in previous situations that resembled the one at hand; he then selects whatever mode of treatment had the most successful outcome in the past. *Id.* at 22.

In ordinary clinical treatment, the purpose is not to gain new knowledge but to repeat a success of the past. *Id.* at 23.

Second, the subject matter and conditions of study are different. "In laboratory work, the experimental material is an intact animal, a part of a person or of an animal, or an inanimate system; in clinical treatment, the material is an intact human being." *Id.* at 22. The hard scientist initiates the experiment at a time of his own convenience and chooses the material usually without regard to its desire or consent for participation. *Id.* In clinical medicine, the patient initiates the treatment, choosing the time, place, duration, and clinician. *Id.* "The physician is not studying the properties of chemical compounds in a test tube; he cannot postpone dealing with cancer in a patient for fifty years because he hopes by then to have a much clearer insight into the nature of the disorder." *Id.*

Finally, clinical medicine and hard science have markedly different methodologies. A clinician observes at least three types of data for each patient who undergoes treatment: A disease in morphologic, chemical, microbiologic, physiologic, or other impersonal terms; the host in whom the disease occurs and his environmental background, including his personal properties (such as age, race, sex, and education) and external surroundings (such as geographic location, occupation, and financial and social status) before the disease began; and the illness that occurs in the interaction between the disease and its environmental host, consisting of clinical phenomena: the host's subjective sensations, or "symptoms," and "signs," which are findings discerned objectively during the physical examination. Feinstein, at 24–25.

Using these data, the clinician determines a present diagnosis (which gives the disease a name and tells what is wrong), a past etiology and pathogenesis (or how it got that way), and a future prognosis and therapy (or what to do about it). *Id.* at 25. Some of the data used by the clinician can often be obtained by examining the patient's fluids, cells, tissues, excreta, roentgenograms, graphic tracings, and other derivative substances. The patient's personal environmental data can often be elicited by nurses, secretaries, social workers, or other interviewers. But the history-taking, physical examination, and the determination of symptoms and signs can properly be done only by a doctor skilled in the clinical procedures described above. *Id.* "Moreover, the [clinical physician's] capacity to make judgments in cases of a kind which he has never seen before must depend ultimately on a cultivated capacity to see equivalences between quite disparate things, that is, on analogy." Murphy, at 9.

In sum, hard or Newtonian scientific knowledge does not comprehend all subjects that theoretically might be subjected to its methodology. It is knowledge of a particular and limited kind, gathered or tested by a particular and characteristic method. T.H. Savory, *The Language of Science* (1953). Although clinical medicine utilizes parts of some hard sciences, clinical medicine and many of its subsidiary fields are not hard sciences. The purposes, criteria, values and methods of hard or Newtonian science and clinical medicine are far from identical. Fred A. Mettler, *The Medical Sourcebook* xxxiv (1959). Consequently, the *Daubert* factors, which are hard scientific methods selected from the body of hard scientific knowledge and methodology generally are not appropriate for use in assessing the relevance and reliability of clinical medical testimony. Instead, the trial court as gatekeeper should determine whether the doctor's proposed testimony as a clinical physician is soundly grounded in the principles and methodology of his field of clinical medicine.

McCleskey v. Kemp

753 F.2d 877 (11th Cir. 1985), *aff'd*, 481 U.S. 279 (1987)

[McCleskey was convicted of murder and sentenced to death in the Georgia state courts. In a federal habeas corpus petition, he challenged the conviction and sentence on several grounds. In support of a constitutional argument, he offered a now-famous *regression analysis* (a statistical technique that is discussed in Chapter 3, Statistical Inference; the regression analysis in this case is also discussed there) that purported to show that the odds of receiving the death sentence were substantially higher for black than white defendants, especially when the victim was white. The petition was rejected at every level.]

Social Science Research Evidence

To some extent a broad issue before this Court concerns the role that social science is to have in judicial decisionmaking. Social science is a broad-based field consisting of many specialized discipline areas, such as psychology, anthropology, economics, political science, history, and sociology. *Cf.* Sperlich, *Social Science Evidence and the Courts: Reaching Beyond the Advisory Process*, 63 Judicature 280, 283 n. 14 (1980). Research . . . is conducted under both laboratory controlled situations and uncontrolled conditions, such as real life observational situations, throughout the disciplines. The broad objectives for social science research are to better understand mankind and its institutions in order to more effectively plan, predict, modify, and enhance society's and the individual's circumstances. Social science as a *nonexact* science is always mindful that its research is dealing with highly complex behavioral patterns and institutions that exist in a highly technical society. At best, this research "models" and "reflects" society and provides society with trends and information for broad-based generalizations. The researcher's intent is to use the conclusions from research to predict, plan, describe, explain, understand, or modify. To utilize conclusions from such research to explain the specific intent of a specific behavioral situation goes beyond the legitimate uses for such research. Even when this research is at a high level of exactness, in design and results, social scientists readily admit their steadfast hesitancies to conclude such results can explain specific behavioral actions in a certain situation.

The judiciary is aware of the potential limitations inherent in such research: (1) the imprecise nature of the discipline; (2) the potential inaccuracies in presented data; (3) the potential bias of the researcher; (4) the inherent problems with the methodology; (5) the specialized training needed to assess and utilize the data competently; and (6) the debatability of the appropriateness for courts to use empirical evidence in decisionmaking. *Cf.* Henry, *Introduction: A Journey into the Future — The Role of Empirical Evidence in Developing Labor Law*, 1981 U. Ill. L. Rev. 1, 4; Sperlich, 63 Judicature at 283 n. 14.

Historically, beginning with "Louis Brandeis' use of empirical evidence before the Supreme Court . . . persuasive social science evidence has been presented to the courts." Forst, Rhodes & Wellford, *Sentencing and Social Science: Research for the Formulation of Federal Guidelines*, 7 Hofstra L. Rev. 355 (1979). *See Muller v. Oregon*, 208 U.S. 412, 28 S. Ct. 324, 52 L. Ed. 551 (1908); *Brown v. Board of Education*, 347 U.S. 483, 74 S. Ct. 686, 98 L. Ed. 873 (1954). The Brandeis brief presented social facts as corroborative in the judicial decisionmaking process. The Brandeis brief "is a well-known technique for asking the court to take judicial notice of social facts." Sperlich, 63 Judicature at 280, 285 n. 31. "It does not solve the problem of how to bring valid scientific materials to the attention of the court. . . . Brandeis did not argue that the data were valid, only that they existed. . . . The main contribution . . . was to make extra-legal data readily available to the court." *Id.*

This Court has taken a position that social science research does play a role in judicial decisionmaking in certain situations, even in light of the limitations of such research. Statistics have been used primarily in cases addressing discrimination.

Statistical analysis is useful only to show facts. In evidentiary terms, statistical studies based on correlation are circumstantial evidence. They are not direct evidence. *Teamsters v. United States*, 431 U.S. 324, 340, 97 S. Ct. 1843, 1856, 52 L. Ed. 2d 396 (1977). Statistical studies do not purport to state what the law is in a given situation. The law is applied to the facts as revealed by the research.

In this case the realities examined, based on a certain set of facts reduced to data, were the descriptive characteristics and numbers of persons being sentenced to death in Georgia. Such studies reveal, as circumstantial evidence through their study analyses and results, possible, or probable, relationships that may exist in the realities studied. The usefulness of statistics obviously depends upon what is attempted to be proved by them. If disparate impact is sought to be proved, statistics are more useful than if the causes of that impact must be proved. Where intent and motivation must be proved, the statistics have even less utility. This Court has said in discrimination cases, however, "that while statistics alone usually cannot establish intentional discrimination, under certain limited circumstances they might." *Spencer v. Zant,* 715 F.2d 1562, 1581 (11th Cir. 1983), *on pet. for reh'g and for reh'g en banc,* 729 F.2d 1293 (11th Cir. 1984). These limited circumstances are where the statistical evidence of racially disproportionate impact is so strong as to permit no inference other than that the results are the product of a racially discriminatory intent or purpose. . . .

Much has been written about the relationship of law and the social science. "If social science cannot produce the required answers, and it probably cannot, its use is likely to continue to lead to a disjointed incrementalism." Daniels, *Social Science and Death Penalty Cases,* 1 Law & Pol'y Q. 336, 367 (1979). "Social science can probably make its greatest contribution to legal theory by investigating the causal forces behind judicial, legislative and administrative decisionmaking and by probing the general effects of such decisions." Nagel, *Law and the Social Sciences: What Can Social Science Contribute?*, 356 A.B.A.J. 356, 357–58 (1965).

With these observations, this Court accepts social science research for what the social scientist should claim for it. As in all circumstantial evidence cases, the inferences to be drawn from the statistics are for the fact finder, but the statistics are accepted to show the circumstances.

NOTES AND QUESTIONS

1. In *Moore*, the court confronted the problem of how to categorize the diagnostic opinions of clinical physicians ("differential diagnosis," as it is termed in medicine) under *Daubert*. Most people would probably describe the practice of Western allopathic medicine as "scientific," and the doctors who conduct it as informed by scientific research, even though they may not be active researchers themselves. The dilemma for the courts is that an individual diagnosis does not—literally, at least—possess *any* of the four identifying features of positivist science that *Daubert* identified. It is not testable: If the single patient fails to recover, that does not "falsify" the diagnosis, no more than the patient's recovery would validate it; most patients get better regardless of what the doctor does. Individual diagnoses are almost never published. An individual diagnosis has no "error rate." And although differential diagnosis is itself generally accepted in the medical community, that label cannot be meaningfully applied to a single case. Courts are thus left with two choices: to reject differential diagnosis as failed or "bad" positivist science, or to accept it as a "good" example of a hybrid: a complex, judgmental art that is nonetheless strongly rooted in a vast body of evolving scientific knowledge. In the opinion above, the Fifth Circuit opted for the