2007

The Irrelevance, and Central Relevance, of the Boundary between Science and Non-Science in the Evaluation of Expert Witness Reliability

D. Michael Risinger

Follow this and additional works at: https://digitalcommons.law.villanova.edu/vlr

Part of the Evidence Commons

Recommended Citation
Available at: https://digitalcommons.law.villanova.edu/vlr/vol52/iss4/1

This Symposia is brought to you for free and open access by Villanova University Charles Widger School of Law Digital Repository. It has been accepted for inclusion in Villanova Law Review by an authorized editor of Villanova University Charles Widger School of Law Digital Repository.
THE IRRELEVANCE, AND CENTRAL RELEVANCE, OF THE BOUNDARY BETWEEN SCIENCE AND NON-SCIENCE IN THE EVALUATION OF EXPERT WITNESS RELIABILITY.

D. MICHAEL RISINGER*

I.
Kick at the rock, Sam Johnson, break your bones:
But cloudy, cloudy, is the stuff of stones.

II.
We milk the cow of the world, and as we do
We whisper in her ear, “You are not true.”

RICHARD Wilbur’s little two-stanza poem is entitled “Epistemology,” and it nicely encapsulates the problem of knowledge. Extreme anti-realist and relativist positions are not so easy to refute, but on the other hand, actually believing them in any operational sense is, almost literally, insane.

Caudill and LaRue’s No Magic Wand is a breath of fresh air in the difficult three-way relationship between “science studies,” science, and the law. By rejecting the radical constructionist element of science studies that has historically been its most public face, Caudill and LaRue have

* Professor of Law, Seton Hall University School of Law. I would like to thank Larry Laudan for helpful comments and suggestions on the text, and Lesley Chenoweth Risinger for insisting on clarity, analytically and editorially. The faults and errors, alas, remain my own.

1. My copy is in THE NEW POETS OF ENGLAND AND AMERICA 325 (Donald Hall, Robert Pack & Louis Simpson eds., 1957), a source that pretty well dates me too.


3. I have followed Caudill and LaRue’s lead in using the umbrella designation “science studies” to cover those schools of approach to the sociology of science variously referred to as “science studies,” “science and technology studies,” “sociology of scientific knowledge,” and “social studies of science.” The first three designations are referred to at CAUDILL & LAARUE, supra note 2, at xv, and the latter phrase is the name of the leading journal in the field. The umbrella organization for the area is called the Society for the Social Studies of Science (SSSS).
cleared the way for science studies and its insights to play a constructive role in the way law regards science.4

I have just spoken of the difficult three-way relationship between science studies, science, and the law. The existence of any relationship at all may come as a surprise to most lawyers. I suspect that most lawyers, perhaps even most legal academics, or even evidence scholars, have never even heard of “science studies” as a specific discipline, or if they have, they have only the vaguest notion of what it is.5 I must admit that I was in the same situation until just a few years ago. I had heard rumors of something called the “science wars” involving critiques and defenses of the objectivity of science. I formed the impression that the critiques were heavily influenced by post-modernism, an approach to thought which generally makes me break out in hives.6 With that in mind, I gave the science wars little thought.

4. “Science is a product of both (i) observation and experiment, with respect to natural reality, and (ii) norms, conventions, and expectations within the scientific community.” Id. “Oddly, however, while social constructivists do not idealize science, they do idealize the social aspects of science to a degree that the successes of science are either ignored or eclipsed.” Id. at 3.

5. Thoroughly unscientific (assuming this designation has any meaning) surveys undertaken by me have tended to confirm this. Enrolled members of the science studies guild have been known to suspect the same thing. See Michael Lynch & Simon Cole, Science and Technology Studies on Trial: Dilemmas of Expertise, 35 Soc. Stud. Sci. 269, 277 (2005).

6. By “post-modernism” I mean the hyper-relativist approach to knowledge utilizing “deconstruction” as a methodology, under the influence of Jacques Derrida, Michel Foucault, and many others. See generally PAUL R. GROSS & NORMAN LEVITT, HIGHER SUPERSTITION—THE ACADEMIC LEFT AND ITS QUARRELS WITH SCIENCE 71–106 (1994). This school of thought, with its constant vamps on radical skepticism, infected many parts of the academy like a plague from the 1970s through the 1990s, most particularly literary studies, history, some forms of social science, and, to a certain degree, law, where critical legal studies in its worst excesses strangely coexisted with the often equally extreme versions of law and economics perpetrated by the right. Both had (and to a certain extent, continue to have) a self-indulgent obscurantist style of exposition. The law and economics obscurantism is often realized by feckless mathematization under circumstances where no real numbers with any empirical value are ever likely to be available. The obscurantism of Critical Legal Studies has been manifested through the adoption of the syntactically tortured, jargon-ridden and unnecessarily verbose style of what is asserted to be standard English that has marked much of post-modern writing in many areas (the law not the least). These excesses of near-unintelligibility have always seemed to me to be driven by three forces: the desire for a shared mysterious rhetoric to act as a badge of membership in the tribe (in essence a shibboleth); an arrogance (on the part of some) reflecting the opinion that one is so brilliant that first drafts will do, and that clarity is the job of the exegetes to come; and, alternatively or concomitantly, a fear of what actual clarity of exposition might reveal.

Such texts often remind me of the joke about the yeshiva student who went to hear a traveling rebbe. Upon his return he was asked what he thought of the rebbe, and he said, “He’s a very smart rebbe, and I understood what he said.” The next week, the student again went to hear a different traveling rebbe. When he returned he was again asked what he thought, and he replied, “He’s a truly brilliant rebbe. I only understood half of what he said.” The next week the student
My first formal introduction to science studies was probably in 2000, when I joined David Caudill, among others, at a symposium at Washington and Lee University on the Daubert trilogy and its implications. The contribution to that symposium by Caudill and his co-author Richard Redding was a lengthy article entitled Junk Philosophy of Science?: The Paradox of Expertise and Interdisciplinarity in Federal Courts. It was from this article that I first got an actual glimpse of the world of science studies and its constructionist bent. I also discovered that its main target seemed to be an extreme positivist view of science, in which the product of science was entirely the function of external reality somehow channeled wholly through the scientist as medium. I further learned that a main repository of this view in the legal literature was the West treatise Modern Scientific Evidence, to which I had contributed two chapters. So perhaps my role here is supposed to represent the wandless wizards of science implied by the book’s title and that of the symposium. If so, I will do my best both to disappoint and not to disappoint such expectations.

I must admit that even after the 2000 conference I was still not too inclined to expend a lot of time exploring the mysteries of science studies. From what I somewhat apologetically admit was a rather cursory reading, I still associated it with the dread “post-modernism.” But I kept meeting more people like David, whom I liked and respected and who seemed to think there was something “there” in science studies, chief among them Jennifer Mnookin and Simon Cole. So I did a little investigation. Probably again went to hear yet another traveling rebe, and when asked what he thought he declared enthusiastically, “He’s the most brilliant rebe in the world. I didn’t understand anything he said.”


8. Caudill and Redding seem to adopt, at least as an analytic organizational tool, Bruno Latour’s four-category characterization of available positions on the continuum between pure positivism and pure constructionism: (1) “radical,” allowing no role for external reality; (2) “progressivist,” regarding science as dominantly constructed but conceding that some exterior reality leaks in at various points; (3) “conservative,” regarding scientific knowledge as dominantly the product of exterior reality while conceding some role for social influences; and (4) “reactionary,” regarding proper products of proper science as fully objective, determined by the exterior world. See id. at 726–86, citing Bruno Latour, One More Turn After the Social Turn, in THE SCIENTIFIC STUDIES READER 279 (Mario Biagioli ed.) (1999). Latour claims to reject all these (though for what reason he does not say), excoriating most particularly “the marsh of wishy washy scholars who add a little bit of nature to a little bit of society and shun the two extremes.” Though I have never particularly aspired to be wishy washy, I am one of those middle grounders, as the reader will see in due course. Apparently, so are Caudill and LaRue. See CAUDILL & LaRUE, supra note 2, at 4.


---

2007] THE BOUNDARY BETWEEN SCIENCE AND NON-SCIENCE 681
bly not enough to actually justify what I am going to say, but this is a sym-
posium and I have the conch, so I am going to say it anyway.

It seems to me that science studies in its various manifestations have
suffered from a self-inflicted wound. It has made itself easy to associate
with the worst excesses of post-modernism as post-modernism played out
across various academic disciplines. In the main, perhaps science studies
practitioners did not go so far in that direction as their critics later
charged, at least not consistently. But if they were not card carrying post-
modernists, they were at least sometimes opportunistic fellow travelers,
and there was still plenty of excessive rhetoric to choose from, collections
of which have been made in various places by Larry Laudan,10 Susan
Haack,11 and Alan Sokal and Jean Bricmont,12 among others. Defenders
of science studies sometimes respond that such quotations do not re-
present the field or that they were taken out of context.13 However, there
has rarely been any clear repudiation of that rhetoric.

The main bones of contention concern a constellation of intercon-
nected issues that really are not, or do not have to be, central to science
studies as a discipline,14 but seem to have become central to its identity:
First, the existence vel non of any reality external to the senses of human
observers; second, the role of such external reality (if any) in the products
of the enterprise of science; third, whether there is anything special about
the methods of science as a human activity; and fourth, whether science
can be properly characterized as marked by progress. It may be helpful to
set out two polar extremes of response to these questions, which we will
call the heroic positivist response and the extreme post-modernist
response.

10. See Larry Laudan, Science and Relativism 55 n.7 (1990) (quoting David
Bloor and Barry Barnes); id. at 74 (characterizing positions of Bloor and Harry
Collins); id. at 106 (quoting Bloor and Barnes's definition of knowledge as "any
collectively accepted system of beliefs").

11. Susan Haack, Defending Science—Within Reason 21 (2003) (quoting,
among other things, Harry Collins's infamous statement that "the natural world
has a small or non-existent role in the construction of scientific knowledge")
There are numerous other quotations from Barnes, Bloor, Collins, Bruno Latour
and Steve Woolgar throughout the book, most particularly in chapter 7, "A Modest
Proposal—The Sensible Program in Sociology of Science" and its accompanying
notes. Id. at 179—205.

12. See Alan Sokal & Jean Bricmont, Science and the Sociology of Science: Beyond
War and Peace, in The One Culture?—A Conversation About Science 27-28 (Jay
A. Labinger & Harry Collins eds.) (quoting Collins, Latour, Barnes and Bloor, and
Stanley Aronowitz). See generally Alan Sokal & Jean Bricmont, Fashionable Non-

13. See, e.g., Michael Lynch, Is a Science Peace Process Necessary? in The One Cul-
ture?—A Conversation About Science, supra note 12, at 50–52. See also Harry M.
Collins & Robert Evans, The Third Wave of Science Studies: Studies of Expertise and

14. See Susan Haack's "sensible program" for the sociology of science in
Haack, supra note 11, at 194–201.
The heroic positivist\textsuperscript{15} would say that there is external reality, that science has a special methodology that, when properly applied, allows a practitioner to discover new details of the contours of that external reality and the regularities that define it; that it is the external world of reality that accounts for and is reflected in the products of science (and thus they are "objective"); and that science is both cumulative and progressive, in that old knowledge is built on and built up toward an understanding that is better than yesterday (and not as good as tomorrow), leading ultimately to a perhaps distant but realizable Laplacian\textsuperscript{16} full account of the world accessible to the senses.

The extreme post-modern position would be that there may be no external reality; that even if there is, that science has no special way of getting at it; that the products of science are dominantly social constructions to which external reality (if any) contributes very little if anything; and that since the products of science are only valid in and for the human communities that construct them, they are neither objective nor progressive. Science, in short, is just another human game like the construction of mythology in ancient Greece. Ultimately, everything reduces to (unscientific) psychology, politics and power.

Critics of science studies (often, but not exclusively, practicing scientists) accuse science studies people of largely buying into the postmodernist viewpoint.\textsuperscript{17} Science studies people deny this and accuse

\textsuperscript{15} I have at another place characterized this view as “pathological positivism.” See D. Michael Risinger, \textit{Boxes in Boxes: Julian Barnes, Conan Doyle, Sherlock Holmes and the Edalji Case}, 4 \textit{Int’l Comment. on Evidence}, Iss. 2, Art. 3 at 9, available at http://www.bepress.com/ice/vol4/iss2/art3. I refer to it as “heroic positivism” here to emphasize its conceptual reliance on an image of a single heroic thinker as the operative unit in science. It is an extreme version of what Professor Haack refers to as the “Old Deferentialist” view of science. See Haack, supra note 11, at 31–51. It seems to correspond closely with Latour’s “Reactionary.” See Latour, supra note 8.

\textsuperscript{16} The main contribution of Pierre-Simon Laplace (ironically the greatest probability theorist of the early 19th-century) to nineteenth century positivism was the notion that if one attained sufficiently detailed knowledge of the facts concerning a phenomenon, notions of probability and prediction would disappear because all of the past and the future would be entailed in the present knowledge.

We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes.

\textsuperscript{17} See Gross & Levitt, supra note 6, at 42–70; see also Steven Weinberg, \textit{Facing Up—Science and Its Cultural Adversaries} (2001). Gross is a biologist, Levitt is a mathematician; Weinberg (and the previously referenced Sokal and Bricmont, supra note 12) are theoretical physicists. Laudan, supra note 10, is a philosopher of science and Haack, supra note 11, is an epistemologist. For an excellent, detailed and carefully critical intellectual history of the post-WW II tech-
the practicing scientists of merely trying to defend a version of heroic positivism that the various examinations undertaken by science studies people have rendered untenable.\textsuperscript{18} They further say that their relativism is not "cognitive" or "radical" or "philosophical" relativism, but merely "methodological" relativism—an approach to description and analysis of scientific disputes that puts aside any judgment based upon the ultimate epistemic merits of the positions of the various sides.\textsuperscript{19} Only in this way, they say, can one avoid post-hoc "heroic" or "whig" history,\textsuperscript{20} where the contributions of social factors to the performances of only the losing side are up for discussion (e.g., "why in the world did those losers take that position in the face of all this evidence"). The critics of science studies respond that methodological relativism was a late-arrived-at stop-gap cover for a more fundamental relativism, and in addition, that even methodological relativism makes it impossible to examine properly how much the final outcome of the process of science is in fact the product of the phenomenon under examination, and not merely an exercise in social construction like the generation of mythic origin stories in traditional cultures.\textsuperscript{21}

There is evidence that the two sides in the "science wars" are finding some common ground. As it turns out (though this should hardly be surprising), there are very few if any practicing scientists of any repute (and fewer philosophers of science)\textsuperscript{22} who subscribe to an heroic positivist view of science, with its emphasis on a model of individual inquiry and its strong assumption that ultimate truth is, at least in theory, directly know-

\begin{flushright}
\begin{itemize}
\item \textsuperscript{18} See Trevor Pinch, Does Science Studies Undermine Science?—Wittgenstein, Turing and Polanyi as Precursors for Science Studies and the Science Wars, in \textsc{The One Culture}?: A Conversation About Science, supra note 12, at 21-22.
\item \textsuperscript{19} See Harry Collins, One More Round with Relativism, in \textsc{The One Culture}?: A Conversation About Science, supra note 12, at 186-90.
\item \textsuperscript{20} Wikipedia describes "whig" history as: [A] pejorative name given to a view of history that interprets history as a story of Teleological progress toward the present. The phrase was coined by the British historian Herbert Butterfield in 1931, in his small but influential book \textit{The Whig Interpretation of History}. It takes its name from the British Whigs, advocates of the power of Parliament, who opposed the Tories, advocates of the power of the King and the aristocracy. The term has been applied in historical disciplines outside of British history (the history of science, for example) to criticize any goal-directed, hero-based, and trans-historical narratives. See Wikipedia.com, Whig History, http://en.wikipedia.org/wiki/Whig_history (last visited Apr. 6, 2007). Though it must be approached with some caution, I am in awe at the general success of the Wikipedia project, which I would not have myself predicted.
\item \textsuperscript{21} See Sokal & Bricmont, supra note 12, at 38-43. Jean Bricmont & Alan Sokal, Reply to Our Critics, in \textsc{The One Culture}?: A Conversation About Science, supra note 12, at 243-49.
\item \textsuperscript{22} See Laudan, supra note 10, at 149; see also Haack, supra note 11, at 22.
\end{itemize}
\end{flushright}
This heroic positivist view (which I refer to as the “Sherlock Holmes” notion of science), was never particularly tenable, but it did appeal to a certain Victorian sensibility. It died among thoughtful scientists (to the extent it ever lived among them) in the early 20th century with the coming of relativity theory, quantum mechanics, the Heisenberg principle, probabilistic models of multiple causation and expansion of the size of teams required to do serious research in many areas. (It was a scientist, after all, J.B.S. Haldane, who gave rise to “Haldane’s Law,” that the world is not only stranger than we know, it is stranger than we can know.) Everyone of any reflection and familiarity with actually “doing science” will concede that normal science practice includes much that is social, and much that is an exercise in “construction.” Indeed, these themes were present in the pre-“science studies” sociology of science represented by Robert K. Merton, a figure science studies people often seem to look down upon as representing a somehow outmoded “First Wave.”

None of this means that the “Sherlock Holmes” vision of heroic positivism doesn’t live on. It certainly does, and to a very great extent, among

---

23. See Risinger, supra note 15. This is hardly the place to attempt an intellectual history of 19th-century positivism. Suffice it to say that its single most influential voice was Auguste Comte, see Zammito supra note 17, at 6-8, and that it was marked by the notion that all knowledge was reducible to factual propositions accessible to science, that existence in all its manifestations was marked by progress, and that expert practitioners of a generalized “scientific” approach in virtually all spheres of knowledge would discover perfect truths that would in turn dictate social goals and means leading to progress. See Nicola Abbagnano, Positivism, in 6 The Encyclopedia of Philosophy 414 (Paul Edwards ed., 1967). It is easy to see how such a set of notions would appeal in an age of colonialist expansion and industrial progress. It was the dominant popular notion of science among the literate in the formative period in which the law encountered “forensic science,” and in a fairly unreflective way it continues to be a common though distorting lens through which a wide range of people view “science.”

24. In fact, heroic positivism began dying among serious thinkers about science during the same period it seemed to embed itself in popular culture (apparently because of the dramatically visible triumphs of science, especially applied science). For an introduction to the history of the shift to probabilism and fallibilism among scientists in the 19th century, see Larry Laudan, Science and Hypothesis (1981).

25. “My own suspicion is that the Universe is not only queerer than we suppose, but queerer than we can suppose.” J.B.S. Haldane, Possible Worlds (1927), quoted in Columbia World of Quotations. Consider also the following from Lewis Thomas: “The only solid piece of scientific truth about which I feel totally confident is that we are profoundly ignorant about nature. . . . It is this sudden confrontation with the depth and scope of ignorance that represents the most significant contribution of twentieth-century science to the human intellect.” Lewis Thomas, The Medusa and the Snail 73-74 (1973).

26. “We are all, in a commonsensical way, cultural constructionists in our view of science.” Gross & Levitt, supra note 6, at 43 (describing “weak form” of cultural constructionism).


28. See Collins & Evans, supra note 15, at 239. Collins and Evans dispose of the “first wave” (which they assert had “crashed on the shore” by early 1970s) in a single paragraph without mentioning Merton.
the public at large, and in the legal community also, lawyers and judges alike. 29 And as Caudill and LaRue properly point out, it can have pernicious consequences sometimes, 30 but the science community and the science studies community ought to be allies in attacking these.

Science studies proponents, on the other hand, seem to be (sometimes somewhat grudgingly) admitting that an assumption of external reality is (in reality) the only game in town and that science has a strong claim that it is progressive in some sense, mainly springing from its own obviously increased predictive power over the decades, and from the related increased success of applied science in controlling aspects of the external world. 31 So the continued debate seems to be, in the long run, mainly over the proper proportion to be accorded to social construction factors in comparison to the external phenomena under examination, and, in the short run, over the usefulness of “methodological relativism.”

Perhaps we can bring this into some focus by examining a model Science Studies case study as an illustration. I have selected my friend Simon Cole’s work, beginning with his analysis of the petroleum origin debate, for it was in the “science studies” framework that Professor Cole was trained at Cornell and his article “Which Came First, the Fossil or the Fuel?” 32 is an excellent example of the kind of study this approach can generate at its best. Professor Cole noticed/discovered a classic circumstance of the kind that Science Studies feeds on: There is a generally accepted theoretical account of a physical phenomenon (in this case, that petroleum products are the product of geological forces transforming fossil remains). 33 This account is generated by and made the subject of further research by a sub-category of scientists (geologists, specifically, petroleum geologists). A brilliant maverick from outside the proprietor group (in this case Thomas Gold, a person with credentials that makes him difficult to ignore) 34 challenges the theory with a different theory that may be just as good and in some ways superior from the perspective of the claimed general principles of science (data coverage, parsimony, etc.). (In this case, the alternative is that all or most oil and gas are the product of “out-gassing” from the

29. The success of television shows like the various “CSI” iterations shows that heroic positivist views of science are alive and well among the masses.
33. Id. This hypothesis is called biogenesis. The opposing hypothesis is abiogenesis (explaining the origin of oil, gas and coal solely through physical and chemical processes within the Earth).
34. Thomas Gold is a member of the National Academy of Sciences and a Fellow of the Royal Society, with significant contributions to his credit in a variety of physics and astronomy subdisciplines. Id. at 736–37.
Earth’s mantle). Controversy ensues, coupled with various predictable and “unscientific” human reactions. In the end, however, the article concludes where science studies articles are expected to conclude—the alternative hypotheses appear to be equally tenable (or perhaps the maverick is a bit favored rhetorically as the underdog), the hegemony of the standard theory remains unshaken, textbooks stay the same and science in this instantiation looks more like small town politics than a lordly engine of sure and objective truth. Professor Cole maintains a coolly objective stance to the end. There is perhaps a bit of obligatory constructivist rhetoric, but there are no explicit screeds making “strong programme” points. The facts as set out are generally allowed to speak for themselves, but the subject selected makes the “science studies” point by itself sufficiently for it to be selected for summary and comment by that guru of science studies, Harry Collins and his co-author Trevor Pinch in The Golem at Large—What You Should Know About Technology.

Notice a number of things. First, the controversy, in general terms, concerns an important but small piece of geological science. Second, it is examined over the course of a relatively short period of time. No doubt numerous such stories could be told in selected decades in regard to continental drift, for instance. This does not mean that one should conclude that the theory of plate tectonics has not given us a better understanding of the phenomenon of continental drift (vel non) than was the case half a century ago. Progress emerges out of individual messy controversies over time. I am not claiming such progress is inevitable in the individual case. In regard to the petroleum origins debate, as in any individual case, the controversy may be abandoned or go unresolved for lack of sufficient interest, or lack of funds to do the expensive research that might resolve the issue. But just because not every problem raised is

35. Gold basically resuscitated the theory of abiogenesis, which had been a competing theory through the 19th and early 20th centuries, when it faded in the 1920s. Id. at 756–57.

36. “The Strong Programme” (British spelling intentionally retained) is the name of the strongly relativistic approach to the sociology of science originally set out by David Bloor of the University of Edinburgh in David Bloor, Knowledge and Social Imagery (1976).


38. Cole makes use of the analogy to the tides of opinion concerning continental drift. See Cole, supra note 32, at 737.

39. “Many theories of the past, so far as we can tell, were both genuinely referring and empirically successful, but we are nonetheless loath to regard them as approximately true. Consider, for instance, virtually all those geological theories prior to the 1960s which denied any lateral motion to the continents. Such theories were, by any standard, highly successful (and apparently referential); but would anyone be prepared to say that their constituent theoretical claims—committed as they were to laterally stable continents—are almost true?” Larry Laudan, SCIENCE AND VALUES 123 (1984) (commenting on the problems of realist correspondence or approximate correspondence theories of truth in science).
solved and not every issue is subject to progressive resolution, does not mean that there are not other controversies, perhaps even the majority, that lead to progressive resolution.

Third, Professor Cole maintains strict "methodological relativism" in this article. He makes no authorial observations concerning the relative merits of the competing theories. In a study of an ongoing scientific controversy with qualified participants and no obvious external interests involved, this is completely defensible, but such a stance can easily reinforce the notion that evidence (and the external reality that lies behind evidence) is irrelevant to the ultimate resolution of controversies, even while the stance aids understanding the realities of scientific practice on the ground in the individual case studied.

And so Professor Cole's career might have proceeded, had he not encountered fingerprints. I think it likely, when he first encountered the odd status of the claims to "scientific" knowledge that lay behind fingerprint identification, that he hoped and trusted it would be another circumstance where warring claims to the mantle of science would be interesting to explore in themselves, and would result in the end in the predicted (and desired) science studies denouement—the controversy is unresolved because there is no more to be said on the one side than on the other, although the "establishment" view prevails in the end for political or proprietary reasons that can be met with a sophisticated shrug. It is a good thing for all of us that in the end it did not work out that way.

Professor Cole first discovered fingerprints as long ago as 1995, and it was the subject of a number of conference presentations that, judging by their titles, appeared to assume that the claims of the fingerprint identification community, while of course challengeable, were at least as valid as any other claims to "scientific" knowledge. Nonetheless, there appears to have been something in the critical reflection necessary for the shaping of his magnum opus, the prizewinning book Suspect Identities, that caused the growing suspicion evident in sections of that book that the claims of practitioners of fingerprint identification to scientific knowledge—or knowledge of any sort in regard to their more extreme claims—was insupportable not merely in the same sense that the science studies "strong programme" may assume that all claims are ultimately insupportable, but insupportable in a more fundamental and important way.

40. See supra note 19 and accompanying text.
42. Simon A. Cole, Suspect Identities (2001). Suspect Identities is a beautifully written book, which among other things instantly replaced all previous works as the standard history of fingerprint identification in particular and forensic identification in general.
2007] THE BOUNDARY BETWEEN SCIENCE AND NON-SCIENCE 689

From that point on, Professor Cole’s scholarship has become increasingly critical in the best sense, pointing out the weaknesses in the standard accounts of why we should accept the validity of crime scene fingerprint identification, and exploding the myth of infallibility that has surrounded such identification for nearly a century. Consider the following:

The chief problem with this conceptual framework is that it offers no opportunity for assessment. Fingerprinting has constructed a perfect rhetorical system, in which the actual accuracy of the technique is irrelevant. Consider this: a forensic technique for which no measurement is made of examiner accuracy; no match standards are set; little effective regulation is in place; little scrutiny is exercised by other actors in the criminal justice system; all expertise resides in law enforcement; is legally untouchable; and is popularly regarded as “science.” Such a technique might be highly accurate, or it might not be. It actually makes little difference, and we have no way of knowing which state of affairs is in place. Even if fingerprint identification itself is highly accurate and fingerprint examiners are highly ethical, other techniques and other analysts can adopt the same framework with equal facility. The framework functions just as well for inaccurate techniques as it does for accurate techniques. Within this conceptual framework, we have no way of sorting accurate techniques from the inaccurate, or competent analysts from incompetent, other than the occasional forensic scandal or miscarriage of justice, which is the tip of the iceberg that tells us we have a lurking problem.

When I suggest applying the overheated political term “junk science” to forensic fingerprint identification, I mean something different than what one might assume at first glance. By “junk science” I do not necessarily imply that forensic fingerprint examiners routinely produce incorrect results or miscarriages of justice. Certainly, they do produce these things; but how much, we can only guess. Rather, I suggest that the conceptual framework in which fingerprint examiners operate is “junk science.” In particular, the way in which fingerprint examiners vouch for their knowledge is “junk.” Fingerprint examiners may or may not be highly accurate, but in either case backing up testimony by reference to “one hundred years of experience,” “the authority of forty judges,” “one hundred years of adversarial testing,” or “professional judgment based on training and experience” is, to put it colloquially, “junk science.” It is not thoughtful, it is not rigorous, and it is not—whatever the term may ultimately mean—scientific.43

Not much "methodological relativism" here. But instead of going to the other extreme, and demanding the kind of statistically determinate data that characterizes DNA identification before we accept the validity of any identification by fingerprint matching, Professor Cole has concentrated on the real problems, which (conceding that there may be enough randomness in fingerprint patterns to be confident of a ten-print match from two sets of rolled prints) include inter alia: first, the insupportable concept of "zero methodological error rate"; second, the existence of more errors than the myths of infallibility would suggest; third, the necessity of approaching error rates, not globally, but specifically for the potentially more difficult subtasks that result when latent prints are small in extent and unclear in detail; fourth, the fact that no work has been done making it clear when the threshold of reliability is reached as to the extent and clarity of the available latent print go down to the point commonly encountered with crime scene prints; and fifth, the use of case experience as self-justifying information of accurate practice. Professor Cole has thus moved beyond "methodological relativism," whatever its uses may sometimes be, with a realization that when claims invoke the mantle of science without meeting the standards of science, they are not just different approaches to knowledge of the material world—they are often inferior in their epistemic status. Science (whatever it is) is really onto something when it comes to valid and reliable approaches to truth about the material world. It is, in Professor Haack's memorable terms, not "privileged" but "distinguished."

None of this is to say that every fact of interest to the law is the subject of scientific knowledge. This is especially true in regard to trials, as opposed to the factual components of policy judgments that lie behind lawmaking exercises. The facts relevant to policy are closer to the theoretic generalities which are the main hoped-for product of science per se. The facts of a particular case are often of the everyday type about which no particular expertise is needed to inquire into them, and even if expertise

45. See id. at 996–1023.
46. See id. at n.23.
47. See id.; see also Simon A. Cole, Does "Yes" Really Mean Yes?: The Attempt to Close Debate on the Admissibility of Fingerprint Testimony, 45 JURIMETRICS J. 449, 463 (2005) ("Courts have yet to grapple with the lack of scientific studies testing latent print examiners' ability to do what they claim to be able to do.").
49. Often, not always. Not all claims to reliable knowledge about the material world invoke the mantle of science, and many may be of high epistemic quality. See, e.g., infra note 69 (discussing value of "mere empiricks"); see also infra notes 51, 78, 85 and 121.
50. HAACK, supra note 11, at 94–95.
was needed, no science properly conceived (about which more later, of course) would have anything useful to say, at least currently.\textsuperscript{51}

I have just used the words "science" and "expertise," and here we come to the point where I must begin to deliver on the first part of the rather paradoxical title to this article, the part that says that the demarcation problem (the problem of drawing a reasonably determinate line between science and non-science as modes of inquiry and knowledge generation, and thus between knowledge claims that are the product of science and those that are not) is irrelevant in developing approaches for evaluating the reliability of expertise proffered in court. What I mean by this is that the line between science and non-science is initially not particularly helpful in analyzing or organizing the phenomenon of claimed expertise as it exists in the legal setting. Such analysis and organization should come before addressing any questions about where "scientific" expertise fits in, or how it should be handled in comparison to other claims of expertise. Indeed, it is my contention that the label "science," with its inevitable honorific connotations, simply gets in the way in regard to the enterprise of developing a defensible taxonomy of expertise in the legal setting.

I have made some detailed observations concerning such a taxonomy,\textsuperscript{52} and indeed, those thoughts in their current iteration were what I placed in the materials distributed to the continuing legal education participants in this symposium. I will not reproduce those \textit{in toto} here, but a summary is necessary to make my point about the irrelevance of "science" to the undertaking.

My first claim is that one can only begin to get a handle on the notion of expertise by looking carefully at what is concededly non-expert testimony—the testimony of ordinary fact witnesses. What we want from a fact witness is testimony about concrete details concerning the episode that gave rise to the controversy before the court, what we can profitably call "adjudicative facts." Fact witnesses are defined by their interchangeability with jurors—the only informational advantage we can ascribe to a fact witness over a juror is that the witness was someplace at a time the juror was

\textsuperscript{51} Not all rational empirical inquiry is part of science in the modern sense. There was, of course, a pre-modern use of the term which was applied to any organized attempt at knowledge. This usage lives on in such terms as "military science" and "library science." Professor Haack quite rightly regards science as continuous with good practices of empirical inquiry in other areas from history to crime detection to everyday life, an approach she labels "critical common-sensism." \textit{Haack, supra} note 11, at 23–29. Indeed, the products of such inquiry may be as well warranted in their sphere as the product of science narrowly defined, which is a good thing, since bringing real science to bear on many issues of importance to the law would be practically impossible.

not, and thus had direct perceptual access to information that the juror
did not have. Notionally, if their positions of time and space were re-
versed, the juror would become a proper witness and the witness a proper
juror. In this sense they are interchangeable. Expertise exists when a wit-
ness has some advantage over the average juror more than merely the ad-
vantage of time and place in regard to a factual detail of the episode
under examination. But what might such an advantage be that will allow
witness assertions beyond what is allowable for a fact witness?

Here, I adhere to a great functional divide that is sometimes recog-
nized in the literature, the divide between education of the jury as an ex-
pert function, and translation of information as an expert function. When
a witness performs the first (educational) expert function, the witness is
informing the jury of general information that the average juror could not
be expected to know. In general, the witness who educates is also summa-
rizing a great deal of foundational information for the efficient education
of the jury, which is why I have also styled such witnesses "summariza-
tional" experts.\textsuperscript{53} The information provided from such “expertise” must
be rationally usable by the jury to supplement, qualify or correct such labor-
atory of life generalities about the workings of the world that otherwise
comprise the “common sense” that the jurors are supposed to use in evalu-
ating the meaning and implications of the “adjudicative fact” evidence of
ordinary fact witnesses and other sources of information produced at trial.

When a witness performs the second (translational) expert function,
the witness is taking raw information and translating it from its raw form
to a form more usable by the jury. In terms I find congenial, the expert
takes potentially relevant information with little or no usable relevance
from the perspective of the ordinary person (the juror), and translates the
information into a usably relevant form using what the witness claims is a
reliable system of translation.

The fundamental problem with proffered expertise is that of ensuring
that every expert assertion carries with it a proper rational belief warrant.
It is this that is at the heart of any defensible notion of the judge’s role as
gatekeeper. Different forms of expertise are amenable to different forms
of belief warrant and demand different approaches to what constitutes a
proper belief warrant. In addition, different functional contexts may ap-
propriately be said to allow weaker belief warrants or demand stronger
belief warrants. In the case of educational expert testimony, we may use-

\textsuperscript{53} I have generally used the term “summarizational” to emphasize the way in
which the witness processes the information that lies behind the testimony. On
those rare occasions when such witnesses are explicitly recognized, they are usually
referred to as “educational” experts, and that is likely to remain the common term.
\textit{See generally} Ronald J. Allen & Joseph S. Miller, \textit{The Common Law Theory of Experts:
Deference or Education?}, 87 Nw. U. L. Rev. 1151 (1993); Ronald J. Allen & Joseph S.
Miller, \textit{The Expert as Educator: Enhancing the Rationality of Verdicts in Child Sex Abuse
fully divide such testimony again into two domains, everyday summarization and technical summarization.

Everyday summarizational expertise is the testimony we label "expert" that is closest to fact witness testimony. The example I usually use involves a case under the Uniform Commercial Code in which the proper construction of a contract turns on industry practice in the wholesale shoe business:

A witness is called who has been in the wholesale shoe business for thirty-five years. He will universally be declared an expert and his testimony will generally be characterized as "expert testimony." However, to the extent he merely recounts the contours of practice in the shoe industry as he knows them to be from his experience, how does his testimony differ from that of the ordinary fact witness? He is not necessarily possessed of any relevant skills or talents beyond the members of the jury. There is no reason to believe that if any of them had been where he has been and perceived what he has perceived, that they would not in the ordinary course know what he knows. He seems to have nothing more than a space-time advantage over the jurors, no different than a fact witness.

There is an important difference, however, in the desired testimonial function of the ordinary fact witness and that of the shoe man. We want the ordinary fact witness to traffic in empirical specifics. In addition, we want those specifics expressed in the most concrete fashion practicable. When fact witnesses begin to express themselves in more inferential terms, summarizing a number of specific percepts with an umbrella inference such as the word "drunk" or the word "angry," fights start to break out in the courtroom over the propriety of the terms in which the witness is testifying. We need not tarry at length over the unhelpful terms in which those battles are often waged. Suffice it to say that the general principles which should guide the judge in controlling the manner of expression of fact witnesses are reasonably agreed upon: (1) No inferences beyond the capacity of the witness; (2) Even if inferentially conclusory testimony is within ordinary capacity to accurately render, require more concrete testimony and leave the inference to the jury unless; (a) the inference is based upon subliminal percepts not fairly reproduced in testimony; or (b) unless trying to explain to the witness what is desired will confuse the witness and result in a net loss of dependable relevant information.

In the case of the shoe man, such an approach would defeat the whole purpose of his testimony. He is there to give precisely the kind of summary we would not want from the ordinary fact witness, in order to educate the jury about the practices of the
shoe industry as efficiently as possible consistent with giving accurate information. This summary is derived from a data base of many particular and concrete observations over a long period of time. Much of this is now beyond specific recall, but we assume that, as a person who remained long in business, his resultant impressions and conclusions are in general accurately weighted conclusions based on the totality of his experience. If we were to require him to testify in more concrete terms, his testimony would become a series of anecdotes which would not necessarily represent a proper sample of his whole experience. We want him to perform what can profitably be labeled the "summarizational function," and because ordinary fact witnesses are debarred from it, the price of admission for the shoe man is to label him something other than a fact witness. Traditionally, there is only one other label available, so he is declared an expert, which is taken to authorize the summarizational form of his testimony.

To this point, our model of the summarizational expert has been our shoe man, and there has been an assumption that most or all of the knowledge which goes into his summary testimony is knowledge derived from his direct personal experience. This may be the case in a given situation, and it is this direct personal experience summarizational expert function which is conceptually closest to the fact witness function. But in reality, a real shoe man might derive much of his information about the workings of the industry from secondary sources, such as industry meetings, networks of friends, and so forth. He may also have read industry publications of various sorts. In this case the shoe man's testimony will be a summary result not only of his direct personal experience, but of these secondary hearsay sources as well. We could hardly do otherwise and allow him to testify at all, because he himself could not say with confidence which parts of his knowledge were based on personal experience and which on secondary information. But we assume that his first-hand experience has enabled him in various ways to evaluate and internalize the secondary information, with reasonable reliability inhering in the resultant summary. Our tolerance for this hearsay element, and our reliance on the witness as filter of it, is a second way in which a summarizational expert differs from the ordinary fact witness.54

Everyday summarizational expertise offered to educate the jury is perhaps the easiest expertise for the jury to deal with because it is closest to fact witness testimony. Belief warrants spring merely from substantial relevant experience and the plausibility of the knowledge claim. In the shoe

54. Risinger, Functional Taxonomy, supra note 52, at 90–91 (footnote omitted).
man's case, his ability to stay in business in a competitive field is the main circumstance that warrants a belief in the general accuracy of his summary of the relevant contours and usages of the shoe business. This is not to say there are not difficulties to be faced even in regard to such everyday summarizational expertise. The worst difficulty is patrolling the line between the kinds of general assertions that such experience warrants and assertions that are so particularized that they cross the line into specific adjudicative fact assertions. The particular area that illustrates this best, perhaps because it commonly presents the most problems is what I have styled "m.o./argot" testimony by police officers.

It is not uncommon for the prosecution in a criminal case to believe that the jury needs to be educated about general contours of criminal enterprises and ways of operating in order to properly evaluate the adjudicative fact evidence. It is probable that the average juror has no idea how a "badger game" works, what quantities of drugs are normally purchased by end users (and therefore what quantities indicate amounts probably intended for sale rather than personal use), or what the current meaning of the word "spliff" is on the street at a particular time and place. If such knowledge is relevant to the interpretation of the testimony or other evidence in the case, such education seems warranted, and it is not unreasonable to believe that some law enforcement officers would have obtained reliable information concerning those items during their careers. Nonetheless, such testimony can easily slide from education to unjustified adjudicative fact assertion, morphing subtly from "more than an ounce is rarely bought for personal use" through "more than an ounce is a sale amount" to "the amount in this case was a sale amount." Similarly testimony that "the word 'product' is often used to mean a contraband drug" can morph through "the word 'product' usually means 'cocaine'" to "when the defendants referred to 'product' on the tape in this case, he meant 'cocaine.'" Policing the proper boundary becomes even harder when the court is facing one of the detectives who actually investigated the case being tried, since that detective usually has actual adjudicative fact information that he can legitimately testify to as a fact witness. For this reason, judges should in general not allow case detectives to serve also as "educational" experts.55

In addition, such witnesses sometimes try to dress their summarizational expertise up as some sort of translational expertise, often with some sort of claim to scientific status, in order to claim both broader latitude and more weight for their testimony. Offender profilers and crime

"linkage analysts" spring to mind. Obviously the warrant necessary for such assertions ought to be different (and more onerous) than for mere everyday summarizational/educational expertise.

We may now leave our consideration of everyday educational or summarizational expertise and move on to a more controversial form of educational expertise, which we may call academic or technical summarization. Perhaps the clearest example is the testimony of defense experts in criminal cases regarding the weaknesses of eyewitness identification who give no conclusions about the accuracy or inaccuracy of any identifications in the individual case. Rather, these defense experts summarize the relevant findings of research on how various conditions affect the likely accuracy of identifications made under such conditions. Much of the prosecution-proffered "syndrome" expertise falls into the same category: "I don’t know if the witness in this case is making up the charge of rape, but I will say you shouldn’t be too influenced in making that judgment because of the delay in complaining, because research tends to show that a surprisingly high percentage of people actually raped do not immediately complain."

While everyday summarizational/educational experts (the shoe man) have been used without much actual reflection time out of mind, the kind of technical or academic summarizational/educational expertise we are now considering is a late-comer to the courtroom. Many courts have had a difficult time accepting the propriety of such an expert witness who does not give a specific "conclusion" or "opinion." The propriety of such an expert function, however, was specifically recognized in the original advisory committee notes to Federal Rule of Evidence 702, and the great balance of authority now accepts it as at least notionally acceptable.

How to deal with rational belief warrants for such witnesses is not entirely clear. One could argue that such exercises of expertise require less demanding belief warrants than some other kinds of expert testimony because the witness is not making case-specific conclusions. But often the purport of the witness’s testimony applied to the individual case seems so clear that this variable alone does not seem like a proper basis for loosening otherwise proper standards. So perhaps the best thing to say is that such technical educational expert testimony should be subject to the same standards as translational expertise in the same context. What those stan-

---

56. See Risinger, Three Card Monte, supra note 55, at 210-70 (examining history, theory and available validation data for "offender profiling" and crime "linkage analysis").

57. "Most of the literature assumes that experts testify only in the form of opinions. The assumption is logically unfounded. The rule accordingly recognizes that an expert on the stand may give a dissertation or exposition of scientific or other principles relevant to the case, leaving the trier of fact to apply them to the facts." Fed. R. Evid. 702 advisory committee's note (1975). Accord Amended Fed. R. Evid. 702 advisory committee's note (2000). ("The amendment does not alter the venerable practice of using expert testimony to educate the factfinder on general principles.").
standards should be generally cannot be adequately addressed without introducing the notion of science, so I will put that question off until I have dealt with the kind of expert testimony that is most people's very model of expert testimony, translational expertise.

When most people think of an expert performing an expert function in court, they think of the expert testifying to an "opinion" or "conclusion"—some non-obvious factual state that the expert has deduced from material from which the jury could not derive the same conclusion. In performing this function, the expert is claiming to be able to translate information already present in the material into an accurate correspondence with some other non-obvious fact that the jury is not capable of inferring on its own. The language translator is the most basic illustration of this kind of expertise, taking sounds encoded with information, existing in the air of the courtroom equally available (as sounds) to both the jury and the translator, and changing the information encoded in the sounds from a form not rationally usable by the jury to a form that is rationally usable by the jury. All experts who give "opinions" or "conclusions" perform an exactly analogous function.

There are many potential bases on which a witness can rest a claim to be a reliable and valid translator. I have attempted elsewhere to give an account of the logically possible claims to possession of a reliable and valid translation system,\(^58\) and I will not repeat here all that I said there. It is important to note, however, that some translation systems, when stripped of the rhetoric with which they are often clothed, are almost exclusively a claimed subjective skill. They often derive from experience and are personal to the individual translator while other systems can claim objective criteria that can be applied relatively independently of individual skill in many or most circumstances.

The general principle when dealing with translational expertise is to first identify the characteristics of the claimed translation system as it applies to the task presented by the case at hand, and then identify the reasons for believing that the claims of the system to accurate translation are good enough for the legal context in which they are operating.

In order to make this more clear, I will of necessity have to descend to specifics. But before proceeding to examples I would like to point out that, so far, nothing in particular has been said about separating science from non-science. So far, not only has it been unnecessary, it would have been positively obstructive to the enterprise of laying a foundation for the various types of functional expertise we might encounter in court. In my opinion, nothing has retarded proper analysis in this area, in both the courts and the academy, like too much invocation of "science" too soon in the analysis. But later, when we discuss how to approach belief warrants about the translational systems lying behind various claims of expertise, we may no longer be able to avoid examining the phenomenon of

---

claimed scientific knowledge and wrestling with the demarcation problem to boot. Up to now the notion of "science" has been largely irrelevant, but it may become importantly relevant later.

Before proceeding to the renewed importance of the boundary problem, however, I would like to give an example that illustrates that proper standards of belief warrant do not necessarily turn on resolving the science/non-science boundary, and definitely cannot be viewed productively according to some globally applicable uniform threshold of reliability, however defined, above which expertise is admissible, and below which it is inadmissible. Consider land valuation experts (real estate appraisers).

Assume a civil case involving a claim for damages turns on the establishment of the market value of a piece of real estate at a particular time. The very notion of market value is fraught with conceptual difficulty because it is not really a fact, nor a predictive counterfactual (predicting what would have happened in a sale in the real world), but rather a predictive counterfactual under conditions not often obtained in the real world (willing buyer, willing seller, equality and completeness of information, etc.). Nevertheless, the law quite commonly incorporates this strange abstract concept into remedial formulas.

Now market value is not a "fact," but no one claims that it is unrelated to actual facts in the world. Sales of similar items in the real world are as close as we can come to evidence of "market value," and in circumstances where the real world approaches the conditions of the notional concept of the ideal market fairly closely (fungible goods, active high volume market and free flow of information—think money markets) the problems of proof are generally inconsequential. When the actual conditions of real world sales do not approach those conditions very closely (non-fungible goods, relatively low-volume market, non-transportable goods creating geographic sub-markets, incomplete information, often unevenly distributed between buyers and sellers in unpredictable patterns), however, proof of "market value" is inherently imprecise. It is unlikely that we would trust a jury to construct such a market value itself out of the raw materials of evidence of various sales merely produced by the parties.

In this kind of case, the law finds itself in a dilemma. If the remedy for a conceded wrong is made to turn on establishing market value, and if under common circumstances the inherent indeterminacy of the concept would seem to lead only to proof that was "speculative," the law has built the common defeat of justice for failure of proof into the very fabric of its remedial definition. What to do? Allow the plaintiff to dissolve the problem by pouring the universal solvent of "expertise" on it. As we will see, this approach utilizes expertise of very low epistemic quality and turns out to be a brilliant solution to a particular problem, making a silk purse out of a couple of sow's ears. In such a case, both the plaintiff and the defendant will call witnesses they claim to be possessed of "land valuation expertise."
It should be no surprise that the reliability of such expertise by almost any standard of evaluation is low. Such land appraisers are not even in possession of a reliable taxonomic scheme for determining comparability of sales, which is why the list of comparable sales cited by two such witnesses is rarely the same, and often does not even overlap to any great degree.\textsuperscript{59} They generally make no claims to be practicing “science” (though if they did there are a few science studies scholars who would contend that such a claim was irresolvable, and as good as any other claim). Yet, given the issue (not a binary decision about a fact in the world, but an issue literally without a determinate single right answer, even in theory), the applicable burden of persuasion (preponderance), the jury’s ability to actually understand and evaluate the plausibility of the testimony\textsuperscript{60} and the failure of the law that would result if proof were found impossible in most cases both in practice and in theory, the resulting accommodation is brilliant. Plaintiffs’ experts cheat as high on the value as their notion of comparable sales and the “straight face” requirement will allow. Defendants’ experts do the same in the other direction, their “translational” opinion testimony sets a floor and a ceiling, and whatever the jury picks in the range so defined is a legally and morally proper result. Yet I would hope that few readers would think it proper for a criminal defendant to be convicted largely on the basis of prosecution-proffered expertise offered on the issue of his identity as the perpetrator of the charged crime, which expertise was of no greater reliability than the expertise of the average land appraiser.\textsuperscript{61}

What this illustrates to me, and I would hope to everyone, is the total impropriety of any attempt to set up a single “reliability” scale with a single test for the measure of reliability to be reached for admissibility, even if such a scale could be credibly generated. Such a scale would \textit{a fortiori} and by definition either require that we exclude the land appraisers or that we admit against criminal defendants any prosecution-proffered expertise which meets the reliability of land appraisers. This just cannot be right.

\textsuperscript{59} I have been involved in a few cases involving such expertise, and I was initially shocked to discover this lack of overlap. Nonetheless, I know of no published formal study to cite.

\textsuperscript{60} This is generally the case in testimony concerning residential real estate. The variables in play are so near to jury common knowledge that one might be tempted to regard the real function of the appraiser as merely a non-expert summarizing the voluminous material of the comparable sales by analogy to Fed. R. Evid. 1006, except for the fact that the witness is always allowed to go beyond this, both in selection of comparable sales and in giving an “opinion.” None of this is to say that some appraisal situations do not fairly involve variables that quickly get beyond the jury’s everyday competence, nor is it to say that courts should not exercise gatekeeping control when appraisers (real estate or not) cross the line of plausibility. \textit{See}, e.g., Blue Dane Simmental Corp. v. Am. Simmental Ass’n, 178 F.3d 1035 (8th Cir. 1999).

\textsuperscript{61} The land valuation example is examined in greater detail in Risinger, \textit{Functional Taxonomy}, \textit{supra} note 52, at 105.
One is therefore forced to admit that proper reliability standards for expert admissibility vary with function and context.

This should not be read to mean that a deciding court has some free-floating discretion in the matter. For example, although in federal practice the Supreme Court's opinion in *Kumho Tire v. Carmichael*62 (the third case in the famous *Daubert* trilogy)63 does speak of discretion, leeway and flexibility. I have been at some pains elsewhere to show that the Court meant to provide discretion to select the most appropriate criteria of reliability for a particular claim of expertise, not discretion in any broader sense.64 At any rate, what we are discussing here is how things should work, not how courts have failed to make them work in various circumstances.

So it seems the task for a court acting as gatekeeper in regard to claimed expertise is three-fold: it must first determine the most appropriate criteria to evaluate the strength of the belief warrant for the claimed expertise applied to the task presented in the case at hand (which is the same thing as its reliability as the term is used by *Daubert*), and it must next decide the appropriate decision threshold for that type of case and context. Only then can it properly decide if the reliability of the proffered expertise meets the appropriate threshold of reliability for the law's purposes. It is in making these judgments that courts will often not be able to avoid coming to grips with science and the boundary problem.

The reason that the court will have to come to grips with the boundary between science and non-science in making decisions about belief warrant is that the kinds of evidence warranting belief are different, sometimes in degree and sometimes in kind, when one is clearly dealing with expertise based on science and when one is clearly dealing with an expert claim not based on science. It will therefore be necessary for the court to have some notion of science. In addition, in some jurisdictions, the science/non-science characterization determines the test of admissibil-

---


63. *Daubert v. Merrell Dow Pharmaceutical, Inc.*, 509 U.S. 579 (1993), was the initial case, of course, followed by *General Electric Co. v. Joiner*, 522 U.S. 136 (1997), which held that the “abuse of discretion” standard of review, the norm for evidentiary rulings, applies to rulings excluding proffered expert testimony under Rule 702, but did not address the nature of the trial court's discretion.

64. D. Michael Risinger, *Defining the "Task at Hand": Non-Science Forensic Science after Kumho Tire Co. v. Carmichael*, 57 WASH. & LEE L. REV. 767, 773-77 (2000). Justice Scalia apparently agrees: "The discretion [the Court] endorses—trial court discretion in choosing the manner of testing expert reliability—is not discretion to abandon the gatekeeping function. I think it is worth adding that it is not discretion to perform the function inadequately. Rather, it is discretion to choose among reasonable means. . . ." *Kumho Tire*, 526 U.S. at 158-59 (Scalia, J., concurring).
ity, and in most jurisdictions it is officially relevant to admissibility. So it is not amiss to give some account that will be of service in attacking the boundary problem.

But the boundary between science and non-science is not perhaps the only important boundary to understand. There is also the boundary between so-called “pure science” and “applied science.” As Caudill and LARUE perceptively point out, all science in the courtroom necessarily is applied science in an important sense because the law rarely cares about covering generalities independent of their application to determine facts in the particular case.

Like everything else in this area, there is more here than initially meets the eye. As we will see a bit later, even “pure science” as an enterprise must deal in predictions of unknown fact states as part of hypothesis testing, and the preferred mode of testing is often (though not always) one that controls the world in such a way as to bring about the predicted fact state (experiment). The predictive control of nature is potentially central to hypothesis testing even in the purest science. At least in theory, however, the proper end products of pure science are reliable generalities about the regularities of the facts of nature. Such knowledge can often then be used to control the physical world in ways that humans could not control it before, but such applications (and the issue of whether they should be undertaken even if they can be) are not part of pure science ideally conceptualized. Every practical application of the discoveries of science, from H-bombs to iPods, is part of applied science (engineering and technology).

65. That is, those jurisdictions still following the Frye test that derived from Frye v. United States, 293 F. 1013 (D.C. Cir. 1929), which typically applies only to “scientific evidence” (and usually only “novel” scientific evidence at that). See FAIGMAN ET AL., supra note 9, at 10.

66. See, e.g., Fed. R. Evid. 702 (“If scientific, technical or other specialized knowledge will assist the trier of fact. . . .” (emphasis added)).

67. “Whenever science comes into the courtroom, it comes in not as pure theory, but as applied science, and thus looks much like engineering. In court, one asks questions such as ‘Why did this bridge fall down?’ or ‘Is the blood found at the crime scene the defendant’s blood?’ and so forth.” CAUDILL & LARUE, supra note 2, at 10.


69. But importantly, not all technology is based entirely on scientific knowledge. Much of it continues to be based on the trial-and-error tinkering of the “mere empirick” attempting to solve a particular problem of improved technological performance. The term, as with its antique spelling, was originally used by physicians trained in the theories of Galen as an epithet applied to less theoretical practitioners. The fact that much technology is not the product of science is perhaps less surprising than it might at first seem. Before there was any science, in the modern sense, there was plenty of effective technology and the main generator of such expert knowledge was the process of the mere empirick. See Laudan, supra note 10, at 117. Knowledge gained by a mere empirick is based on empirical evidence and observation, but not reduced to formal data. The effort of the mere empirick is always directed toward the solution of a technical problem of impor-
care much about science except insofar as it can become applied science, and they don’t separate the two domains very clearly. Be that as it may, it is necessary to address reliability problems as they arise in application in order to evaluate the belief warrant in the individual case. So after talking of science per se, we will address questions of application and also questions of belief warrant for expertise making no claim to the mantle of science, or making an unsustainable claim.

What follows is a work in progress. A form of it first appeared in an article co-authored with Mark Denbeaux and Michael Saks in 1998. A significant revision and expansion was included in a paper Professor Saks and I presented in March of 2006 to the annual Coronado Conference of the Institute for Scientific Knowledge and Public Policy. While its various forms have thus appeared in co-authored works, Professors Denbeaux and Saks graciously (and perhaps even gleefully, considering what might be seen in some quarters as the text’s amateur presumptuousness) concede that the text and approach are primarily my responsibility.70

It is sometimes said that accounts of complex subjects come in two forms: too true to be good, and too good to be true. The “too true to be good” variety addresses every problem at a high level of magnification. It
is excellent for specialists, but quickly loses non-specialists, such as judges and lawyers, and does not provide them with what they need to deal with the problems that they face. The "too good to be true" variety simplifies to the point of giving an account so divorced from reality that it may make the audience feel as if they had learned something, but it in fact it may be counterproductive in giving a false picture concealing the real problems the audience needs to understand.\textsuperscript{71} What I have tried to do is give an account of modern science as a social enterprise that is good enough for a novice to understand, and true enough to highlight the important points for judges and lawyers in approaching an evaluation of the products of science, in separating them from non-science claims of knowledge. With any luck, the account will in some sense have succeeded to the extent it seems banal. Most of the necessary qualifications will appear in the footnotes, and qualifications deemed unnecessary for current purposes will be banished altogether.

What I am attempting might be called an effort at a simplified normative theory of science as a human practice.\textsuperscript{72} It is simplified for the reasons just given. It is normative in three senses. First, it is normative in the sense that it tries to specify general criteria that must be met for a practice to qualify as science in the modern sense. Second, it is normative in the sense that it asserts the specialness of the products of the described practices as knowledge claims. Third, it is normative in that it is idealized, hopefully not in the sense that draws the ire of the science studies community, but in another sense. The "idealized science" attacked by science studies and by Caudill and LaRue\textsuperscript{73} is based on the assertion (often implicit) that scientific practice, including the practices of all "true scientists," realizes the ideals of the theory of scientific enquiry, and therefore generate "scientific truths." That notion of "idealized science" is rightfully viewed as pernicious in the light of actual practice that scientific studies have had a large hand in describing as it actually occurs. Indeed, as pointed out above, virtually no one involved in the practice of science seriously disputes this (even though this sense of the "idealized" view has con-

\textsuperscript{71} This would yield something akin to the "beermat expertise"—the kind of shallow knowledge about technical subjects one can read off the back of some beermats in Britain, apparently. In the United States, it is found inside Snapple caps that figures prominently in the redoubtable Harry Collins's latest foray into taxonomizing the notion of expertise. See Harry Collins & Rob Evans, The "Periodic Table" of Expertises (2004), www.cf.uk/soci/expert.

\textsuperscript{72} For a useful discussion of the utility of such normative theories, and the inter-relation among theory, practice and criticism, in the context of an examination of rhetoric, see Thomas B. Farrell, Norms of Rhetorical Culture 3–8 (1993). I think it is fair to note, for what it is worth, that what is set out in this article would be not too far away from the way that many, perhaps most, practicing scientists approach the defining characteristics of modern science. See, e.g., Charles Liu, Not Seeing is Believing, 116 NAT. HIST. 52 (2007) (providing account of science given by Charles Liu, Professor of Astrophysics at City University of New York, in article on dark matter issue in astrophysics).

\textsuperscript{73} See Caudill & LaRue, supra note 2, at 3–4, 64–74.
continued popular influence). But when I say what follows is "idealized" I have in mind a different meaning: That the view sets down normative criteria of science and science practice that exist to promote the truth value of its products, and which practitioners must strive to conform to as part of the internal criteria of what it means to be practicing science, even though everyone knows they will never be perfectly attained. Far from being pernicious, such standards of ideal practice (which of course may change over time as awareness of error-producing dimensions of practice come to light),\textsuperscript{74} though they are often as much aspirational as prescriptive,\textsuperscript{75} perform the function of disciplining the practitioners of the enterprise of science into the practices that define it and account for its successes. And some of those criteria are explicitly social. Even Newton and Einstein needed others to check their work, not merely in the sense of convenience, but in the sense that multiple evaluation adds epistemic strength. Indeed, a main strength of science lies in acculturation to good individual practices,\textsuperscript{76} coupled with multiple evaluations\textsuperscript{77} to procedurally check and hopefully cancel inevitable individual bias.

The kind of knowledge that comprises the realm and object of scientific inquiry is knowledge about the universe accessible to the senses (either directly or as extended by validated sense-enhancing technology).

\textsuperscript{74} For instance, the second half of the 20th century saw the general acceptance of the necessity for masking protocols in any research setting using humans as perceivers, raters or interpreters, in order to cancel "observer effects," a methodological improvement leading to greatly increased epistemic strength of data and conclusions. See D. Michael Risinger, Michael J. Saks, William C. Thompson & Robert Rosenthal, The Daubert/Kumbo Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion, 90 Cal. L. Rev. 1, 6–9 (2002). In this regard, "science is a moving target that the standards for successful scientific theories shift with time. It is not just our view of the universe that shifts, but our view of what kinds of views we should have or can have." Steven Weinberg, Facing Up: Science and Its Cultural Adversaries 84 (2001).

\textsuperscript{75} It is often said that a fundamental precept of statements about obligation is that "ought implies can" in the sense that it is wrong to erect standards of morality that require people to do what they cannot do. See Ronald J. Allen & Brian Leiter, Naturalized Epistemology and the Law of Evidence, 87 Va. L. Rev 1491, 1499 (2001). That does not mean, however, that it is wrong to require people to try to come close to an unattainable ideal. They "ought" to come as close as they can, and they "can" (tautologically) come as close as they can. What is required is perhaps related to Lon Fuller's distinction between the morality of prescription and the morality of aspiration. See Lon L. Fuller, The Morality of Law, ch. I (1965). Parts of the ideals of science practice (those prohibiting fraud and related sins for instance) are clearly prescriptive, but other parts are perhaps best looked upon as aspirational, though the aspiration and some effort toward it may in fact be prescriptive.

\textsuperscript{76} It is therefore not surprising that the values Merton described as shared by scientists as a matter of sociological fact—communalism, universalism, disinterestedness and organized skepticism—are closely related to the normative values that give theoretical support to the knowledge claims generated by science. See Merton, supra note 27.

\textsuperscript{77} By "multiple evaluation" I mean to encompass both a collaborative sense and a critical sense.
Modern science recognizes that the knowledge it generates is probabilistic at its root. Nevertheless, its claims to progress are founded on the assumption that knowledge can approach truth asymptotically. Just like the asymptotic function that generates values that approach an axis infinitely closely without ever actually touching it, it may be impossible to reach the truth about the material world, but it is possible to know when one has gotten nearer along the curve of approach.78

The way science generates knowledge about the world accessible to the senses, the typical structures of knowledge that this enterprise yields and the principles by which such knowledge is taken to be validated are, perhaps unsurprisingly, not completely worked out in a way subject to total agreement either among practitioners of science or among those who make the examination of the enterprise their own subject of inquiry (that is, philosophers, sociologists and historians of science). Nevertheless, for our purposes, the general characteristics of the enterprise may be fairly well limned.

Any general account of the scientific enterprise must give due regard to four central characteristics. The characteristics include: the centrality of data, the necessity of taxonomy, the equal centrality of the imaginative generation of hypotheses (and the integrated systems of hypotheses known as theories) and the equally important necessity of the testing of those hypotheses. Each of these four characteristics feeds off the other three and each contains mysteries of its own.

The report of a single observation of a condition (either by the senses or by technical extension of the senses) can be called a datum in the sense that it is foundational to the scientific enterprise, but a random collection of such observations would not generate data in the sense that modern science uses the term.79 Instead, science requires standardization of the

78. Larry Laudan suggests that one difference between at least one type of realist and one type of pragmatist in the philosophy of science is that the realist believes we can judge ourselves so closely to truth that one is justified in referring to certain propositions as true (or approximately true), while the pragmatist accepts that we can tell when we are closer without being justified in having any firm belief that we are very close at all. See Laudan, supra note 10, at 143–45. Even such a pragmatist, however, would probably not object to the use of the word “true” understood as shorthand for “provisionally true” or “taken as true for, now, for purposes of human choice, subject to defeasance by new evidence or better accounts in the future.” A strict Popperian would probably object to the invocation of the word “true” under any circumstances. See Susan Haack, Trial and Error: The Supreme Court’s Philosophy of Science, 95 J. PUB. HEALTH 66, 67 (Supp. No. 1 2005).

79. Here is the place to make some observations about the oft-repeated notion that all observations are “theory laden.” The source of the modern version of this notion is often traced to Willard van Orman Quine’s 1950 essay, Two Dogmas of Empiricism (though Quine referenced the earlier work of Pierre Duhem, there is a great deal of disagreement about the relationship between Quine’s and Duhem’s views). See Zammito, supra note 17, at 17-25. But the idea that human perceptions and observations are interpretively affected, even prior to conscious awareness, by pre-existing cognitive categories whether culturally derived or hard-wired in the brain goes back at least as far as (of all people) Comte, and is a staple of modern
conditions of such reports and a formal analytical system for their organization. A central condition that must be present is theoretical reproducibility of observation (two observers in the same position could perceive the same thing), with strong favor given to reports of observations that can be practically repeated by multiple observers. Stronger favor should be given to reports that reflect dimensions of an event that can be measured and even stronger favor given to dimensions that can be measured quantitatively with some precision and with demonstrable sensitivity, reliability and validity.

These reports must be organized according to some analytic system of similarities and differences that generates hierarchies of categories. In short, the enterprise of science requires an explicit, formal and generally well worked out taxonomy. These taxonomic conditions for the collection and organization of observations are necessary conditions for the enterprise of science, but not sufficient in and of themselves to constitute a science. This is not to say that the descriptive organization of observations is not valuable in itself. Whole branches of human knowledge, including much of history and philosophy, are marked by taxonomic work of this sort. The taxonomic enterprise is not unscientific, but it does not by itself make a science. It is a necessary but not sufficient condition. Science must go beyond descriptive taxonomy of even the most quantified sort to the enterprise of hypothesis and theory creation.

cognitive psychology. "No real observation of any kind of phenomena is possible, except in as far as it is first directed, and finally interpreted, by some theory." THE POSITIVE PHILOSOPHY OF AUGUSTE COMTE 2 (Harriet Martineau trans.) (1855). See, e.g., ULRIC NEISSER, COGNITION AND REALITY: PRINCIPLES AND IMPLICATIONS OF COGNITIVE PSYCHOLOGY 43–44 (1976); Risinger et al., supra note 74, at 12–24. That being said, it seems irrational to believe that every part of every observation is totally determined by pre-existing cognitive investments. The externally existing object generating the phenomenon observed provides something, and like the debate concerning "social constructionism" in general, the issue becomes "how much." Suffice it to say that the term "theory laden" suggests too low a contribution from the object observed, at least under many conditions, and it is part of the job of science to discover the conditions that raise the contribution of the external world and diminish the contribution of such "observer effects." For a further discussion, see supra note 74.

80. What I am setting out here is intentionally loose in regard to the criteria of what can count as formal data. I am personally quite accepting of qualitative observational reports as data that can underlie a scientific enterprise, as long as they are formally recorded according to some standards generally agreed upon in the relevant research community, in an attempt to discipline and standardize the process of observation and recordation.

81. In non-quantitative traditions, each observation may be noted as a part of the data, but in more quantitative traditions, a mathematical aggregation may be referred to as "data."

82. I am not claiming that one cannot have science without perfect taxonomy. That would be self-defeating, since perfect taxonomy is as unattainable as is perfect anything else. The necessary taxonomic conditions are themselves asymptotic to perfection and to be judged pragmatically in regard to the problems to be addressed.
A scientific hypothesis is a statement about inter-relationships between taxonomic items or categories. Strong preference is given to hypotheses formulated in such a way that they can generate predictions about unknown states of fact and as a result be subjected to empirical testing. A theory is a set of interconnected hypotheses of varying generality that account for a wide range of phenomena.\(^83\)

A hypothesis that cannot at least in principle lead to empirically testable predictions is a metaphysical proposition that is by definition not part of science.\(^84\) A hypothesis that is potentially subject to testing but has never been tested is unproven and cannot be treated as a source of dependable scientific knowledge.\(^85\) A hypothesis that has been subjected to

\(^{83}\) This is a good place to mention the "underdetermination thesis," which holds that no theory is uniquely determined by a set of observations, that is, that there are always potential alternative theories covering the observations even though those theories are inconsistent with the first theory under consideration. This is generally seen as another Quine proposition. See Zammito, supra note 17 at 25–33. This is also likely true in some sense, but in its strongest form it is perhaps near-trivial in practice, because as Laudan points out, the practical context always presents problems of choosing between warrants for a theory and any of its known rivals, for which good reasons for preference can exist. See Larry Laudan, Demystifying Underdetermination, in SCIENTIFIC THEORIES 267, 277 (C. Wade Savage ed. 1990). So the fact that theories are not uniquely determined (in theory), assuming this to be the case, is of little import in practical theory choice, and of virtually no use to practicing scientists; it is of as little interest to many of them as debates about radical skepticism among philosophers are to detectives attempting to solve a crime, or lawyers, judges and juries involved in the process of determining the facts giving rise to a particular controversy.

\(^{84}\) One must be cautious not to go too far with this statement. Imaginary non-testable entities are often part of theory construction, though the theory itself is expected to yield testable predictions. And conjectures that do not lead to immediately obvious testable hypotheses may be part of the process of theory building by what C.S. Pierce famously (and still somewhat mysteriously) labeled abduction. See Daniel Dennett, CONSCIOUSNESS EXPLAINED 262–263 (1991).

\(^{85}\) One should be careful not to conclude that hypotheses that have undergone testing according to the standards of science can never be a properly warranted source of reliable knowledge. For a further discussion, see supra notes 51, 69, and infra note 121. Indeed, as Susan Haack has said, "not all, and not only, scientific evidence is reliable." Haack, supra note 78, at 66. For this reason, both Professor Haack (not unlike Professor Laudan) appears skeptical of virtually any approach to the demarcation problem, preferring to get down to the specific reasons for a strong or weak belief warrant for any claim, and simply bypassing the question of ultimate label as unnecessary once this is done. See id. at 69. There is much to recommend this approach. It is a great corrective for confusing the bestowal of the honorific "science" with the necessary presence of reliable knowledge. I believe, however, that a familiarity with a general notion of what constitutes the characteristics and strengths at the core of the enterprise of modern science can assist a court or a lawyer in separating weakly warranted from strongly warranted claims, both within the core of science, and in other areas of empirical inquiry. Professor Haack herself seems to have some such a notion in mind when she suggests that "it seems both feasible and useful to try to ensure that judges understand the more commonly employed scientific ideas they are likely to encounter most frequently." Id. at 71. Here it seems that there must be some general criterion implied for the selection of the ideas that are "scientific." All I hope to have done is flesh out a version of this in a way that might be useful to its
substantial empirical testing and has not been falsified may properly be treated as validated for the time being. Although all empirical testing is essentially a probabilistic enterprise and thus no hypothesis is ever subject to absolute verification, the nature of both the data and the fit of the hypothesis with other more or less validated principles may push the probability of the truth of the hypothesis (and its entailed conclusions) so high that it would be crazy to spend much time worrying about the residual probability of falsity for any practical purpose.\textsuperscript{86}

The fourth major aspect of the scientific enterprise is real-world empirical testing of the truth of the relationships asserted by hypotheses and the theories that both derive from and generate hypotheses. Like the taxonomic aspect, the empirical testing aspect involves careful observations of phenomena through the senses. However, empirical testing is focused on the observations capable of testing (falsifying) a hypothesis, whereas taxonomically (and theoretically) important observations may be made in a much less directed way.\textsuperscript{87} The empirical testing enterprise contains a whole constellation of special problems involving, among other things, how to determine what observations and conditions of observation are in fact consistent or inconsistent with a hypothesis.\textsuperscript{88}

Data collection, taxonomy building, hypothesis making/theory generation and empirical testing are not totally separate enterprises, of course, and every practicing scientist will have familiarity with each process, at

\textsuperscript{86} For instance, under certain conditions DNA analysis may establish that the random match probability that anyone other than the accused in a criminal case was the source of a particular residue is so low that the inference that the defendant was the source is "virtually" or "morally" certain, and worrying about the residual doubt is neither reasonable nor required. But one is required to be sure the conditions are present.

\textsuperscript{87} "The most exciting phrase to hear in science, the one that heralds new discoveries, isn't 'Eureka!' but rather 'Hmm...that's funny.'" The observation is often attributed, probably wrongly, to Isaac Asimov. Of course, "that's funny" suggests an unexpected observation, an anomaly, and not all anomalous observations lead to advances. The role of and tolerance of anomaly in science is one of the most difficult and contested of all subjects in the philosophy of science. It was a centerpiece of Thomas Kuhn, The Structure of Scientific Revolutions, especially ch. 6, Anomaly and the Emergence of Scientific Discoveries. For a further discussion of the role of and tolerance of anomaly in science, see Laudan, supra note 10, at 71-92.

\textsuperscript{88} The special problems also include the implications of new data on theory choice. The exact specification of criteria for proper choice between known competing theories, and the epistemic status of that choice, especially given the possibility of theories not yet conceived, is a matter of substantial debate among philosophers of science. Some are comfortable with treating chosen theories as true until displaced, while others have suggested that the word "true" inevitably overstates the warrant for a theory. See Larry Laudan, Beyond Positivism and Relativism (1996). For further discussion, see supra note 85. The relativists, of course, say that there is never any truth-related reason for theory choice.
least within the confines of their area of interest. Each process feeds on
the other, with new data putting pressure on existing taxonomies, providing
material that hypotheses and theories must take into account, hypotheses and theories providing predictions to which testing must be directed,
and new observations made while testing hypotheses which must then be
absorbed into taxonomies. It is important to note that individual practit-
oners within the enterprise of science may show marked predilections for
one function or another. People whose strength is hypothesis generation/
theory building may, for reasons of esthetics or ego, display quite “un-
scientific” commitment to the validity of their hypotheses in advance of
testing (which may sometimes bear them out and sometimes shatter
them). In fact, it may even be that such an “unscientific” emotional com-
mitment is necessary to drive most such people through the act of creative
imagination necessary for the hypothesis generation. Since science is a
human enterprise involving a large number of individuals, it is to be ex-
pected that mixed or impure motives will drive many individuals and re-
sult in work manifesting a range of vices from wishful interpretations of
data to fudging to outright fraud. Nevertheless, given the range of extra-
science views and interests represented in a scientific community, and
their inevitable conflict, the community basis of scientific practice gener-
ally operates to cancel out many manifestations of bias. In addition, the
socialization into the common standards for the practice of science, the
community emphasis on testing and the existence of professional rewards
for carrying it out, hopefully ensure that, in general, claims do not get too
far ahead of the evidence or stay there for too long.

Uncommonly, a single individual may combine substantial talents in
all four foundational areas. For example, Darwin was a first-class primary
observer and taxonomist as well as theoretician who was quite sensitive to
the need for testing. Others may not. Einstein made few primary observa-
tions and was not particularly interested in the dirty details of confirma-
tory (or disconfirmatory) testing (though he of course recognized the
necessity for it, and involved himself when he felt it necessary). 89 Never-
theless, whatever the idiosyncrasies of the individuals involved, for an en-
terprise and its products to qualify as a science, its group practice must
manifest all four functions in a balanced, ongoing and dynamic way, even
if its individual practitioners do not always do so.

These are the main commonalities of all modern science. Each sepa-
rate area of investigation will have its own strengths and weaknesses and
its own supplementary principles of good practice in the fulfillment of the
four common functions. For instance, science is easiest where the phe-
nomena in question present highly fungible manifestations. If we are in-
vestigating the behavior of atoms of a certain isotope of silver, we are

89. Einstein encouraged Dayton Miller to undertake interferometer studies in
the 1920s that would test Einstein’s relativity theory. See Collins & Pinch, supra
note 37, at 42.
justified in treating all such atoms alike for most purposes. This fortunate circumstance eliminates the risk of sampling error and most arguments concerning the external validity of experimental observation. On the other hand, when we deal with biological systems (including humans), individual variation in most phenomena of interest is the rule rather than the exception, making valid generalities harder to establish, more subject to controversy and more likely to be expressed in terms of large set probability rather than more exact prediction. Whatever the area and the special considerations involved, however, there must be a manifest community practice in the four foundational functions. This is the engine and the strength of the scientific enterprise.

To this point, we have dealt with an idealized description of what is often called "pure" science, albeit with a human face. This account follows the traditional notion of the end of pure science: knowledge of fundamental regularities in the material world. Of course, much scientific inquiry is driven at least in part by desires for control of the world as well as knowledge of it, that is, control beyond what it necessary for the core pure science function of testing hypotheses. In reality, much, perhaps most, scientific research is driven by prospects of useful application (and the psychological and perhaps other rewards that accompany it) as well as a pure thirst for knowledge. 90 This has been a theme from the beginnings of the emergence of modern science as a recognizably distinct approach to knowledge, and it is a virtually inevitable by-product of the politics of funding, whether from public or private sources. While the boundary between pure science and applied science is useful conceptually, in practice it is a boundary that is very fuzzy. It is often easier to draw a boundary between scientific and other kinds of knowledge than it is to draw a boundary between pure and applied science. And, as Caudill and LaRue quite properly point out, science in the courtroom is of necessity applied science, since we are only concerned with its implications for the particular case at hand. 91 It is therefore important to understand the special problems of knowledge and controversy that mark the application of sc-

90. "Pure science" was once considered a branch of philosophy (natural philosophy). In this regard, consider the following from William Seeley, a neurologist and neuroscience researcher at the University of California, San Francisco, speaking on recent research dealing with the brain's role in producing the self:

Once we know which brain regions are involved in self-representation, I think we can take an even closer look at which cells in that brain region are important and then look deeper and say which molecules within cells and which genes that govern them lead to this vulnerability [to dementias]. And if we've done that, we've gotten closer to disease mechanisms and cures. That's the best reason to study all this. It's not just to inform philosophers.


91. See CAUDILL & LA RUE, supra note 2, at 10 (noting that science never comes into courtroom as pure theory).
entific knowledge to individual predictions and evaluations, as well as the concomitant control of the world that derives from such application.

First, many technical applications of science-derived knowledge, those that are used to predict and control individual circumstances in the world, involve non-controversial and well established claims of knowledge.\(^\text{92}\) When you think of it, this is not only often as it should be; it is almost as it must be.\(^\text{93}\) Nobody wants a bridge engineer to use controversial bases of prediction in the design of a bridge. Indeed, there is a theory of the sociology of engineering which asserts that the tendency when incorporating new knowledge into design is to over-engineer, to provide both redundancies and large error margins.\(^\text{94}\) It is only later when those initial designs (bridges, for instance) succeed that new designs reducing margins of error are generated, which often move in later designs into zones of high failure risk.\(^\text{95}\)

Pure science in its ideal version is value neutral in regard to all values except its "verististic" or "truth conducive" internal values (sometimes referred to as cognitive values).\(^\text{96}\) This strong claim is actually capable of strong defense.\(^\text{97}\) To the extent goals of control begin to influence agendas of inquiry, however, science is not, as an enterprise, value neutral in its selection of problems to explore. In addition, application of scientific knowledge is never value neutral, because it is directed toward goals of control, and there is always a normative choice underlying any decision concerning what to control, and why. The presence of such value motivations in regard to much that is claimed to be science in the courtroom


This is the element that distinguishes applied science from basic. Surprise is what makes the difference. When you are organized to apply knowledge, set up targets, produce a usable product, you require a high degree of certainty from the outset. All the facts on which you base protocols must be reasonably hard facts with unambiguous meaning. The challenge is to plan the work and organize the workers so that it will come out precisely as predicted. For this, you need centralized authority, elaborately detailed time schedules, and some sort of reward system based on speed and perfection. But most of all you need the intelligible basic facts to begin with, and these must come from basic research. There is no other source. In basic research, everything is just the opposite. What you need at the outset is a high degree of uncertainty; otherwise it isn’t likely to be an important problem. You start with an incomplete roster of facts, characterized by their ambiguity; often the problem consists of discovering the connections between unrelated pieces of information. You must plan experiments on the basis of probability, even bare possibility, rather than certainty.

Id.

\(^\text{94}\) See Petroski, supra note 92, at 595–99.

\(^\text{95}\) See id.

\(^\text{96}\) See Lacey, supra note 68, at 45–66. See generally Laudan, supra note 10.

\(^\text{97}\) See id. at 1–87.
calls for special caution in judging how well the process generating the asserted scientific knowledge has lived up to the ideals of science practice.98

This summary account of the social process of science is not inconsistent with much of anything of any importance said in Daubert or Kumho Tire, and it should actually make it relatively easy to deal with the status of some shaky claims to scientific knowledge put forth in the courtroom. Take Simon Cole’s main subject of study, fingerprint identification. The theory of fingerprint identification is both quite simple, and quite obviously not the product of any process that would be recognized as science in the modern sense. The foundational tenets are, first, that the friction ridges on human skin are the product of some process that introduces a high degree of randomness into their patterns; second, that such patterns remain stable over a person’s life; third, that the pattern is unique to each person; and fourth, that human beings can be taught to compare a print of initially unknown origin to a print of known origin and, at least much of the time, make perfect source attributions with no possibility of error.

As Professor Cole has amply demonstrated, the general theory was generated in the absence of much formal data, and it has never been put to any rigorous formal testing. In addition, its claim of “zero error rate” betrays the 19th century “Sherlock Holmes Science” origins of the theory. No modern theory of application would claim a “zero error rate” for any process, especially anything using a human being as the instrument of evaluation. Therefore, identifications made pursuant to the theory cannot derive whatever epistemic value they have from their derivation according to the standards of science (almost no formal data, no testing—no science). This does not necessarily mean that such identifications, or the theory pursuant to which they are made, have no other sources of supporting evidence, but it does make it inappropriate to regard them as entitled to whatever epistemic credit is due the methods of science (and it makes the fingerprint identification practitioners’ claims to the mantle of science insupportable).

When faced with these virtually incontestable lacunae, the fingerprint identification community responds in the main by saying that long success in practice supplies a proxy for the missing formal data and formal testing. There is occasionally some good sense in this kind of claim, but the problem is that it usually goes too far by ignoring whether there is unambiguous feedback revealing accurate and mistaken results in practice.99 As Professor Denbeaux and I have written:100


100. The following text is from Mark P. Denbeaux & D. Michael Risinger, Kumho Tire and Expert Reliability: How the Question You Ask Gives the Answer You Get, 34 Seton Hall L. Rev. 15, 67-72 (2003). Footnotes have been renumbered to
Putting aside the overblown nineteenth century language of absolute uniqueness in which the claims of fingerprint identification are usually expressed, the main claim may be reformulated in more acceptable modern terminology, thus: Human skin contains a fairly sizable extent of ridged skin on the palms and fingers of the hands and the soles and toes of the feet. The usual theoretical account for its function is that the ridges increase friction on the surfaces, where increased friction for traction and gripping would be of survival benefit. Such work as has been done on the subject indicates that the pattern of ridges in any given individual is constant throughout life. Setting aside the question of whether no two people share the “exact” same pattern of ridges and perceptible detail associated with them, it is clear that not every person has the same pattern of ridges as every other person across the entire extent of their ridged skin. Indeed, though surprisingly little defensible formal research has been done on the question, it seems apparent enough from anecdotal information that variation is so common that “exact matches” across the entire range of ridged skin are vanishingly rare (if they occur at all in the human population now alive). There is apparently some mechanism at work in the fetal development stage which triggers the generation of the ridges by a process that exhibits a fair amount of randomness at a fine level of organization and detail (though the patterns fit general cat-

conform serially to the current article, and in some instances have been supplemented.

101. See United States v. Llera Plaza, 2002 WL 27305, at *2 (E.D. Pa. 2002) (hereinafter Llera Plaza I) (quoting the testimony of Dr. William Babler, President of the American Dermatoglyphics Association, given in United States v. Mitchell, 199 F. Supp. 2d 820 (E.D. Pa. 2002)); see also Cole, supra note 42, at 108. [Note: Llera Plaza I was vacated, superseded, and depublished by United States v. Llera Plaza, 179 F. Supp. 2d 549 (E.D. Pa. 2002) (hereinafter Llera Plaza II)] Though it was initially given the citation 179 F. Supp. 2d 492 in the Federal Supplement advance sheets, it was not published in the bound volume. After an outcry from the academic community that the opinion might be lost to history, Westlaw decided to continue to carry the opinion. Hence the use of the Westlaw references at this remove.


103. Variations in detail are often attributed to random in utero influences, mostly of a mechanical sort. See Cole, supra note 42, at 114. However, more may be going on than that. A start is being made at understanding how various processes, including biological processes, can generate self-organizing patterns displaying such a combination of order and randomness by beginning with a set of relatively simple conditions and subjecting them to fairly simple algorithms which instantiate both positive and negative feedback. See Scott Camazine, Patterns in Nature, Natural History, June 2003, at 34, 40; see generally SCOTT CAMAZINE ET AL., SELF-ORGANIZATION IN BIOLOGICAL SYSTEMS (2001).
egories of pattern at a grosser level of examination). However, because the ridges are generally curvilinear in complex ways, describing the amount of randomness and the likelihood of a random match is a daunting theoretical problem. Even describing what constitutes a match is a problem, because, contrary to popular belief, matches are often not manifested in anything resembling perfect superimposability. The curved and deformable nature of surfaces upon which prints of ridged skin may be left, and the deformable nature of skin itself, often defeats exact superimposition, so that even with prints reflecting large extents of ridges, matching may be an exercise in complex topographical judgment in accounting for such (usually mild but perceptible) deformities preventing superimposition.

So, while the formal research necessary to justify such a statement with formal data has not been done, and the empirical and theoretical work which would give a proper explanatory account of the mechanism behind the organization of ridged skin has not been done, it seems uncontroversial in any serious way to say that the amount of randomness in ridge organization is such that "no two people" share the same pattern in a confusable way across the entire extent of their ridged skin. We do not, however, use the entire extent of ridged skin for identification purposes. Partly because of convenience in "rolling" such prints and partly because it is the print most likely to be left on a surface inadvertently, we use only the ridged skin on the balls of the fingers. The standard practice known to everyone who has ever been "fingerprinted" is to ink the balls of the fingers and roll them onto a card in boxes marked out for each digit. The result is a set of ten prints of known orientation comprising about one square-inch each, for a total of roughly ten square-inches of ridged skin. Again, as in the case of the entire extent of ridged skin, the formal research necessary to establish random match probabilities for two sets of ten prints from different people has not been done, but it seems fair to conclude that such probability is sufficiently minuscule not to trouble the practical certainty which we seek in the law. When an unidentified body is found, and a ten print card is rolled from the fingers of the

104. See Cole, supra note 42, at 115 (defining classes of arches, loops and whorls).
107. Or is just beginning. See references supra note 103.
corpse, and it is found to match one on file with a law enforce-
ment agency, doubts about the belief warrant for the identifica-
tion would seem trivial.

At the other extreme, however, it is clear that there is a lower limit of certainty. If a glass found in a room where a murder had been committed had a smudge on it which showed clearly only one portion of one ridge one sixty-fourth of an inch long (a very "partial" "latent print"), neither its curvature nor any microscopic detail connected with it would allow a confident identification. Since no one knows its orientation, or which digit it came from, it would have to be compared with every short length of every ridge on every print of a candidate card, and no one knows exactly how many such short lengths of a ridge might match it in any randomly selected ten-print card.

So in regard to the admissibility of fingerprint identification, there would seem to be two potential issues of reliability, one trivial and one extremely important. The first would challenge the admissibility of any identification derived from fingerprint comparison, on the ground that without formal research and quantified statistical modeling, its reliability could not be established. Such a challenge is puckish, quixotic, and in some ways constructive, but in others not.109 In the end, it is doomed to failure, and not simply for the wrong reasons either. First, there has been little empirical study that tends to indicate that, at least in regard to large clear areas of ridged skin, variability is large and coincidental matches are at least very rare.110 Second, the extensive use of ten-print comparisons for identification of unknown persons followed by later confirmation of identity from other sources and no known record of error can be said to form a practical accuracy feedback loop unique among forensic identifica-

109. Such positions can be used to paint all critics of forensic science as radical bomb-throwers and extremists deserving of small consideration. For a milder and more nuanced, but still, in our view, much too global version of a similar characterization, see Edward Imwinkelried, Flawed Expert Testimony: Striking the Right Balance in Admissibility Standards, CRIM. JUST., Spring 2003, at 28 (asserting critics demand unreasonable global exclusion, and apparently arguing for (to our minds) unreasonable global admission).

110. The so called "50k x 50k study" testified to by Donald Ziesig of Lockheed Martin Information Systems, which computer-compared each of 50,000 individual rolled loop class prints from white males with each other, was perhaps such a study, see Llera Plaza I, 2002 WL 27305, at *3, but probably any such characterization is wrong. It was really a test of the Automated Print Identification System (AFIS) that assumed the individuality of prints rather than testing that hypothesis. This study is referred to as S. B. Meagher, B. Budowle & D. Ziesig, 50k vs. 50k Comparison Test (1999) (unpublished) in Stoney, Scientific Basis, supra note 98, at 258 n.3, § 34:35, and must be taken with something of a large grain of salt, since it was FBI commissioned and still has never been published. Its many weaknesses are discussed in Cole, supra note 44, at 1046–1047 n.334, and authorities there cited.
tion techniques. While more defensible research is to be encouraged, a global challenge to the reliability of all fingerprint identification is a non-starter.

The second potential challenge is the important one. It is based on the boundary problem described above: identification of a practical threshold of reliability for "partial prints." What standards should be applied to insure that identifications from a small area of print found at a crime scene are sufficiently reliable for purposes of the criminal law? Here, the absence of formal data ought to be more troubling under Daubert and Kumho Tire. This is especially true because fingerprint experts either disagree on how to characterize the threshold of reliability, or more commonly, claim that such a threshold cannot be described at all. This is the result of the addition to fingerprint examination over the last decades of new sources of information (often now collectively referred to under the title "ridgeology") which makes old thresholds fail in some circumstances.

Extensive clear prints such as ten-print cards might be quickly matchable by general pattern at the first general level of observation (sometimes called "the first level of analysis"), and

111. See Scientific Status, supra note 102, at § 34:47, at 362-64 (discussing whether fingerprint procedures have been empirically validated). Professor Cole reports a single case of misattribution in a death case, but that involved the recovery of a single latent thumb print from the badly decomposed corpse after "some skin restoration using tissue builder." See Cole, supra note 44, at 1013.

112. See Scientific Status, supra note 102, at § 34:46, at 360-62; Cole, supra note 42, at 262-63.

113. "Ridgeology" as a term appears traceable to a 1983 pamphlet by David Ashbough, a member of the Royal Canadian Mounted Police, entitled "Ridgeology." The specific detail to which he referred, such as the presence of pores and characteristics of curvature, had been known and considered for some time (the use of pores even has its own term, "poroscopy"), but Ashbough's radical claim that identification was always a gestalt which could never be subject to any threshold system of points (which had been foreshadowed by a resolution of the International Association of Identification, the leading organization of fingerprint examiners) was embraced by many. See Cole, supra note 42, at 261-66.

114. Or rather, which makes old thresholds overly conservative in the eyes of some. See Cole, supra note 42, at 263.

115. Llera Plaza I, 2002 WL 27305, at *4-*5 (Ashbough testimony). This is all part of what is now billed as the "ACE-V" methodology, a "methodology" so lacking in methodological detail as to be, upon reflection, nearly hilarious. The A stands for "assess," that is, look at a latent print and decide whether it is too smudged or small even to try to analyze it, and whether any apparent detail ought to be ignored because it represents a "double tapped" or overlapping print. The C stands for "Comparison," and that means, well, the examiner is to compare the known and the latent print, though there are apparently no fixed standards for performing such a comparison. Rather, it is based "on the training and experience of the examiner." E stands for "evaluation," which means that the examiner decides if the two are similar enough to declare that they are a match, without reference to any particular notion of minimum points of correspondence, and V stands for "validation," which is a non-blind checking of the first examiners work by a second
confirmed by correspondence of individual landmarks (often called minutiae)\textsuperscript{116} at the next level of magnification (often called the "second level of analysis").\textsuperscript{117} These landmarks were the Galton\textsuperscript{118} solution to the curvilinear nature of ridges, identifications of characteristics which could serve as discrete units of analysis, such as the point where one ridge divides into two (often called a "bifurcation"), or the division of two ridges followed by their closure again (a "lake"), etc.\textsuperscript{119} The correspondence of such landmarks, the number of ridges separating them, and the relative direction and distance of their separation, are the stuff of determining the "number of points of comparison" between two prints. However, at yet higher magnification (referred to sometimes as "third level analysis"),\textsuperscript{120} a clear print will show yet more supplementary information detail, including the width and shoreline of individual stretches of ridge, and the presence of pores separated by variable distances. Herein lies the rub. Traditional reliability thresholds typically required from seven to sixteen landmark points of comparison, with no unexplained differences.\textsuperscript{121} Adding the third level of magnification means, according to most examiners, that fewer traditional points are necessary in a clear print because the new detail can make up for fewer landmarks in individual cases.\textsuperscript{122} Why the new details of ridgeology cannot simply be assimilated into the "points of comparison" available to make up a conservative quantified minimum is not completely clear.\textsuperscript{123} Given the subjective

examiner. This is the "scientific technique" which the government in \textit{Llera Plaza} argued "met all four of the \textit{Daubert} guidelines." \textit{Llera Plaza II.} 188 F. Supp. 2d 549, 560 (2002). Judge Pollak found ACE-V not to be "scientific," but appears to have taken it seriously as a "methodology." \textit{Id.} at 561–569.

116. Traditionally (per Galton) the term "minutiae" (singular, "minutia") was synonymous with "Galton points." \textit{See} Stoney, \textit{Scientific Status}, \textit{supra} note 42, at 257–258; Cole, \textit{supra} note 42, at 79–80. There may be a trend toward applying the term to the even smaller "third level" detail. \textit{See} \textit{Llera Plaza I}, 2002 WL 27305 at *2, where, according to Judge Pollak, FBI Fingerprint Unit Chief Stephen Meagher used the term in this way.


118. Sir Francis Galton (1822-1911) was a British biostatistician, geneticist, eugenicist and fingerprint pioneer. \textit{See} Cole, \textit{supra} note 42, at 80 (noting ridge characteristics, also known as minutiae, ridge details, or Galton details).

119. \textit{See} \textit{id.}

120. \textit{See} \textit{id.}

121. \textit{See} \textit{Scientific Status, supra} note 102, at § 34:34, at 352 (noting disagreement concerning number of points to compare).

122. \textit{See} \textit{id.}

123. The argument seems to have two aspects. First, some landmarks, such as a "trifurcation," are so rare that their presence even without much else might be enough for confident identification; second, the process is claimed to be a subjective gestalt process which is not rationally subject to universal thresholds made up of specified criteria. \textit{See} \textit{id.} (same). While the latter may actually describe what examiners do, it would seem desirable to hold them to some sort of statable mini-
nature of the evaluation at the boundary, the necessity, for the purposes of the law, of a mandated threshold in some form would seem most consistent with the policies of Daubert, Kumho Tire, Rule 702, and the standard of proof beyond a reasonable doubt in the criminal law. This is especially true because such evaluations at the boundary are usually performed without any masking protocols to prevent suggestion or expectation from affecting the results, and without any line-up type foils which, in this area, could be easily supplied.\(^{124}\) However, again consistent with Kumho Tire, such determinations of the appropriate reliability threshold should be dealt with in cases which arguably present specific examples of the boundary problem.

One might attempt to counter these claims that the threshold of reliability has not been appropriately established by asserting that the experience of practitioners has shown the threshold’s limits. This is, in fact, the claim that is made for fingerprint identification, and in general it is the claim that lies behind most experience-based translational systems. This is a strong claim under some conditions, but, unfortunately, a weak claim under the circumstances that usually occur in the forensic identification disciplines (fingerprint, handwriting, visual hair comparison, bitemark, toolmark, firearms, etc.).

In some experience-based areas, including those sometimes used by various judges to illustrate the acceptability of non-science, experience-based expertise, such a claim is reasonable. In activities such as beekeeping, harbor piloting, auto repairing or plumbing, the empirics of ordinary every-day practice give unambiguous feedback on correct and incorrect applications.\(^{125}\) However, in the forensic identification disciplines this is

\(^{124}\) See Risinger et al., supra note 74, at 43.

\(^{125}\)Let us continue to examine what kinds of information can yield a proper belief warrant for a claim of experience-based methods or skills beyond self-belief. There appear to be two main sources of such information: practical success and scientific testing of claims. It should not be too surprising that we believe that actual scientific testing of claims is epistemically privileged, and trumps all when it has been done and done properly. However, like epidemiology in the case of risk-increase causation in toxic tort, such testing is expensive and difficult to do across the whole range of claimed practical areas of expertise that are proffered in legal proceedings. In the absence of high-quality testing, are there ever any circumstances that can take its place and provide adequate belief warrants for the purposes of the law?

In regard to claimed expertise at determining specific facts, at any rate, there would seem to be two necessary conditions for such a belief warrant: first, that in the ordinary practice of the claimed methodology or skill, there are objectively unmistakable right and wrong results in most
rarely the case, for reasons I and others have dealt with at length elsewhere. Suffice it to say there is often no objective index of right and wrong identification, only the decision of a jury to believe the testimony of the expert. In such circumstances, the experts’ belief in their own accuracy, or that of the members of the guild to which they belong, is not a sufficient basis for inferring accuracy in fact.

This is where, of course, science comes in again, or should come in. When a non-science claim of experience-based expertise cannot derive sufficient support from the circumstances of its practice, it must be subject to testing according to the appropriate standards for undertaking such tests, which are those adopted by the enterprise of science, standards that have been in part responsible for the demonstrated successes of that enterprise both in theory and in application. Indeed, until such testing is done, it is of questionable rationality to allow testimony of experience-based experts operating on the boundaries, especially when the context is prosecution-proffered identification expertise in criminal cases.

This is not to say that all products of real science are sufficiently reliable for use by the law, either in policy-making contexts or in litigation. Sometimes it is clear that the certainty of results from what appears to be proper science practice is being oversold or distorted, and the admission of such results can easily be attributed to the inappropriate impact of the

---

cases of application, and second, that there is a generally inescapable penalty for wrong results. Under these circumstances, it is at least tenable, at any rate, to believe that humans may develop generally reliable practical methods and skills. Though the practitioners may not be able to give any useful account of the reasons for their success (being only what would have been called in an earlier time “mere empiricks”), if all the law cares about is the success of the methods and skills developed in such circumstances, then the judgments of such cooks (or beekeepers, or chicken sexers) may be proper candidates for admission into evidence.

Of course, clearly apparent right or wrong results, and unambiguous feedback regarding success or failure, are only necessary conditions for a belief warrant about experience-based methods or skills; they are not in themselves always sufficient. There are other conditions which may reinforce or undermine reliability even in the presence of such conditions, and therefore affect both the tenability of belief warrants and the question of admissibility under Rule 702. The emphasis in Kumho Tire is on reliability of the expertise in the circumstances of the case. New Rule 702 requires that proffered expertise be the product "of reliable methods"... [T]here is one enormous reliability-undermining condition which applies to all expertise, but most heavily distorts expertise which is experience-based and relies on human subjective judgment. I refer to so-called "observer effects," particularly those which result from conditions giving rise to expectation and suggestion from which the expert has not been insulated by appropriate masking techniques. In addition, much research indicates that the distortions resulting from such unmasked suggestion and expectancy are reinforced significantly by the kind of team identification and desire to win which are virtually inevitable in the adversary process.

*Id.* (footnotes omitted). Testing of the reliability of skills which has been done in settings without such variables (which would be the norm in the usual design of such tests) cannot establish that the skills survive in the presence of the precursors of such effects.
mystique of science. However, one can easily err in the opposite direction as well. One of the most important of Caudill and LaRue's points is that judges who approach science from the heroic positivist perspective they learned in high school or obtained from the popular media, may expect too much of science, and tend to dismiss scientific knowledge as junk simply because it is uncertain. In addition, such judges may lose sight of the fact that the decision thresholds that are used by scientists to establish consensus may not be the same as those that are proper for the uses to which the law wishes to put the information. In particular, there is a disconnect between the high standards of science and the lower official standards of certainty which are supposed to control allocation of responsibility in civil cases, a point courts have often (though not always) overlooked in spite of its being pointed out repeatedly by the academic community.

So what can we say in general? In general, I believe we can say that what we need in judging the admissibility of proffered expertise is an appropriate belief warrant; that is, good reason to believe that the claimed expertise, given its specific form and the methods and conditions of which it is a product, is appropriately reliable in regard to the case-specific question upon which it is proffered. With this in mind, we can generate a series of questions about how to approach the reliability of expert claims in litigation, independent of the classification of those claims as proceeding from "science" or "non-science," as follows:

1. What is the case-specific target issue to which the expertise is directed?
2. What is the case-specific claim of expertise?
3. What available information bears on a rational belief warrant in regard to the reliability of this specific claim of expertise?
4. What is the appropriate standard of certainty for such a belief warrant given the kind of case, the issue involved, the distribution of the burdens of production and persuasion in the case, and the standard of proof involved in regard to the issue upon which the expertise is proffered?

126. See CAUDILL & LARUE, supra note 2, at 15–30 (discussing whether trial judges are overly skeptical of scientific experts).

In our contribution to the symposium, which forms the subject of much of the discussion by Caudill and LaRue, Professor Denbeaux and I undertook in some detail to illustrate how each question should be addressed in the context of a clear “scientific evidence” proffer (utilizing a hypothetical toxic tort case), and a case involving the proffer of experience-based expertise (the *Kumho Tire* case itself). It will be noted that all of these questions except the third can be answered by the court without reference to the boundary between scientific and non-scientific knowledge. Thus, for both the initial taxonomic examination of the

128. See *Caudill & LaRue*, supra note 2, at 49–83 (discussing and critiquing the various papers presented at a symposium of which I was principle organizer). The symposium was entitled *Expert Admissibility: Keeping Gates, Goals, and Promises* and took place in February 2003. The resulting papers were later published in Vol. 33, No. 4 and Vol. 34, No. 1 of the *Seton Hall Law Review*.

129. See *Denbeaux & Risinger*, supra note 100, at 34–48 (identifying task-specific reliability question for explicit products of science).

130. See *Denbeaux & Risinger*, supra note 100, at 34–48 (identifying task-specific reliability question for explicit products of science).
phenomenon of expertise, and for approaching most of the questions dealing with expert admissibility in an individual case, the science/non-science boundary problem is irrelevant. However, in evaluating the information available on the existence of an appropriate belief warrant, the science/non-science boundary must be addressed, because the nature of evidence for belief warrants in the two domains is different. Thus, as the title to this paper promised, the boundary issue is both irrelevant (to some important issues of expert testimony) and centrally relevant (to the question of how to establish a belief warrant in regard to a particular proffer).