# Lost in Translation: Animal Models, Human Biology, and Expert Evidence Under Rule 702

by Jeremy Beckham*

* J.D., *magna cum laude*, 2026, Lewis & Clark Law School; Certificate in Litigation & Advocacy Skills. Judicial Law Clerk, U.S. District Court for the District of Oregon. The views expressed herein are the author's alone and do not represent the views of the U.S. District Court for the District of Oregon or any of its judges. The author wishes to thank Professor Joanna Perini-Abbott for her invaluable encouragement, advice, and mentorship.

## I. Introduction

A growing body of empirical research has found that the use of animals to model the biology of humans is a profoundly flawed experimental paradigm.[1] Fundamental genetic and biological differences between species lead to a "failure of translation," even in cases where the underlying research is "well-designed and carefully executed."[2] As a result, data derived from animal models have minimal predictive value for claims about human health and can even be misleading. For example, one would erroneously predict that chocolate is toxic in humans based on experiments carried out in dogs.[3] There is increasing recognition of this "translational failure" in the scientific community as well as within those federal agencies that regulate the development and approval of pharmaceuticals or determine funding priorities for the nation's largest biomedical research projects.[4]

This empirical failure of animal models has ramifications in the courtroom. Trial courts have long been required to act as an "active gatekeeper" in making admissibility determinations regarding expert or scientific testimony.[5] Performing this "gatekeeper" role means that judges are "charged with the responsibility of screening unreliable scientific evidence."[6] This judicial "gatekeeping" responsibility was reaffirmed in the 1993 landmark decision *Daubert v. Merrell Dow Pharmaceuticals*, as well as recent amendments to Rule 702 of the Federal Rules of Evidence (FRE).[7] In the context of toxic torts and product liability litigation, courts have

---

[1] *See infra* Part III.
[2] Jessica A. Bolker, *Animal Models in Translational Research: Rosetta Stone or Stumbling Block?,* 39 BIOESSAYS 1700089 (2017).
[3] Fiona Finlay & Simon Guiton, *Chocolate Poisoning*, 331 BMJ 633 (2005).
[4] *See infra* Part IV.
[5] TAL GOLAN, LAWS OF MEN AND LAWS OF NATURE 263 (2004).
[6] *Id*. at 263–264.
[7] Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 585 (1993); FED. R. EVID. 702 advisory committee's note to 2023 amendment.

repeatedly grappled with expert testimony that offers an opinion or states a claim about *human health*, but is based on data derived from *laboratory animal experiments* ("animal model evidence").[8] *Daubert* was itself a toxic torts case supported by such animal model evidence. *Daubert* transformed how federal courts evaluate expert testimony by endorsing a more "flexible" approach.[9] Yet the evidentiary record shows that the plaintiffs' claim that Bendectin caused their birth defects relied largely on animal model evidence and minimized human-based epidemiological evidence. [10] But the *Daubert* plaintiffs' theory was incorrect. The overwhelming weight of the human-based clinical and epidemiological evidence demonstrates that Bendectin does not cause birth defects in humans, and the drug has since returned to the market.[11] The *Daubert* plaintiffs made essentially the same "false positive" error that one would make by using data from dogs to question the safety of chocolate. Thus, *Daubert* should now also be viewed as a cautionary tale for courts when ruling on the admissibility of animal model evidence.

This article argues that courts should reflect the growing consensus of the scientific community and relevant regulatory agencies and view animal model evidence as unreliable. Although Rule 702, as further interpreted by *Daubert* and its progeny, requires a case-by-case analysis, courts can and should adopt a position of "presumptive inadmissibility" for animal model evidence. Given the demonstrated empirical failure of animal models as a scientific

---

[8] Amanda Hungerford, *Back to Basics: Courts' Treatment of Agency Animal Studies After* Daubert, 110 COLUMBIA L. REV. 70, 71 (2010); Kristen Ranges, *Vermin of Proof: Arguments for the Admissibility of Animal Model Studies as Proof of Causation in Toxic Tort Litigation,* 34 GEO. ENV. L. REV. 303, 304 (Winter 2022).
[9] *Daubert*, 509 U.S. at 594.
[10] *Daubert*, 509 U.S. at 583; *see also* Joint Appendix, Affidavit of Frederick Crescitelli, Ph.D., (Dkt. No. 183), 1992 WL 12012184, at *65aa ("[A]nimal studies are also helpful and generally and reasonably relied upon to help determine whether or not a drug is teratogenic . . . Moreover, it is scientifically improper to rely solely upon epidemiology studies to determine whether or not a drug is a teratogen.").
[11] Shelley R. Slaughter et al., *FDA Approval of Doxylamine-Pyridoxine Therapy for Use in Pregnancy,* 370 NEJM 1081, 1082 (2014) ("Aside from the fact that a considerable quantity of data, both direct and indirect, has failed to produce evidence of Bendectin-associated teratogenicity, the withdrawal of Bendectin may actually have had adverse effects on pregnant women.").

paradigm, a position of presumptive inadmissibility is reasonable, practical, and consistent with how courts treat other categories of unreliable pseudoscience, including polygraph ("lie detector") tests.[12] If a "presumptively inadmissible" rule had been adopted earlier, it would have mitigated confusion about other scientific controversies, and averted past failures in the courts' fact-finding process, including the landmark tobacco product liability litigation. A presumptive rule of inadmissibility for animal model evidence is fundamentally consistent with the overarching purpose of the FRE in "ascertaining the truth and securing a just determination."[13]

To evaluate whether courts should treat animal model evidence as presumptively inadmissible, it is necessary to first situate modern evidentiary doctrine within its historical development. The law of expert testimony has long struggled to reconcile the limits of contemporary scientific methodologies with legal decision-making, and the challenges posed by animal model evidence are best understood as part of that broader history.

## II. The Evolution of Expert Testimony and Judicial Gatekeeping

Most foundational legal doctrines concern only the normative or the procedural. They are creatures of human construction and values. But the theory and rules of evidence are conceptually distinct. While evidence law is not designed "for the exhaustive search for cosmic understanding," it is nevertheless concerned with how and when descriptive claims are presented to a fact-finder.[14] Courts use the rules of evidence to ensure that such statements have some indicia of reliability before they are introduced. This helps achieve the purpose of "ascertaining

[12] United States v. Scheffer, 523 U.S. 303, 311 (1998) ("Most states maintain *per se* rules excluding polygraph evidence. . . . Whatever their approach, state and federal courts continue to express doubt about whether such evidence is reliable.").
[13] FED. R. EVID. 102, "Purpose."
[14] *Daubert*, 509 U.S. at 597.

the truth and securing a just determination."[15] Other intellectual pursuits are also chiefly concerned with "ascertaining the truth" by understanding and describing objective reality. Most notably, the pursuit of science. As a result, discussions about evidence law "police the boundaries between disciplines."[16] Nowhere is this boundary blurring more apparent than in the domain of expert testimony, where a person offers testimony based on specialized knowledge beyond the understanding of ordinary lay persons. Thus, the use of expert testimony in courtroom proceedings has long provoked controversies that sit "at the intersection of the two dominant institutions of science and law."[17]

For nearly 500 years, common law courts have engaged in an integrative cross-disciplinary approach to understanding reality and evidence. In other contexts, some have described this general epistemological paradigm of wanting to pursue harmony of sense-making pursuits as "consilience."[18] The touchstone of this integrative approach is the allowance of expert testimony in jury trials in cases where specialized knowledge would assist the factfinder. Justice Saunders, sitting on the King's Bench, offered an explanation for expert testimony rooted in the same precepts of consilience in 1554:

> [I]f matters arise in our law which concern other science or faculties, we commonly apply for the aid of that science or faculty which it concerns. Which is an honourable and commendable thing in our law. For thereby it appears that we do not despise all other sciences but our own, but we approve of them and encourage them as things worthy of commendation.[19]

---

[15] FED. R. EVID. 102, "Purpose."
[16] Jill Lepore, *On Evidence: Proving* Frye *as a Matter of Law, Science, and History*, 124 YALE L.J. 1092, 1098 (2015).
[17] GOLAN, *supra* note 5, at 1.
[18] *See generally,* EDWARD O. WILSON, CONSILIENCE: THE UNITY OF KNOWLEDGE (1998).
[19] Buckley v. Rice Thomas, 75 Eng. Rep. 186, 192 (1554).

The need to "apply for the aid" of expert testimony rose in prominence during the 16th and 17th centuries, as English courts were incrementally transitioning away from a system where juries "inform[ed] themselves" and thus had no use for witnesses at all.[20] Under the old system, jurors were specifically selected because of their personal connection to the facts and case being adjudicated. In its place, courts were transitioning to an "instructional" jury system, where juries were expected to hear the case with no prior knowledge of the matter at hand, in order to fairly and dispassionately receive and weigh evidence offered at trial.[21] Lord Mansfield described the new instructional juror as "white paper," who could consider the issue "merely as an abstract proposition upon the evidence produced before him."[22] But from the earliest cases decided with the "aid of science or faculty," problems emerged. The history of science reveals that superstition, pseudoscience, and magical thinking have long plagued even the most eminent scientific institutions and minds.[23] When these faulty paradigms get transposed on to the legal system, miscarriages of justice occur. An early case of trial testimony by a "person of great knowledge" reveals this problem:

> This was the sum and substance of the Evidence which was given against the prisoners concerning the bewitching of the children before mentioned. . . . There was also Dr. Brown of Norwich, a person of great knowledge, who after this evidence given, and upon view of the three persons in Court, was desired to give his opinion, what he did conceive of them, and he was clearly of opinion, that the persons were bewitched, and said, [t]hat in Denmark there had been lately a great discovery of witches, who used the very same way of afflicting persons, by conveying pins into them, and crooked as these pins were, with needles and nails.[24]

---

[20] Learned Hand, *Historical and Practical Considerations Regarding Expert Testimony,* 15 HARV. L. REV. 40, 44 (1901).
[21] John H. Langbein, Renée Lettow Lerner, & Bruce P. Smith, HISTORY OF THE COMMON LAW: THE DEVELOPMENT OF ANGLO-AMERICAN LEGAL TRADITIONS 238–39 (2009).
[22] Mylock v. Saladine, 1 Bl. W. 480, 96 Eng. Rep. 278 (1764).
[23] *See generally,* THOMAS S. KUHN, THE STRUCTURE OF SCIENTIFIC REVOLUTIONS (1962); ROY PORTER, THE GREATEST BENEFIT TO MANKIND: A MEDICAL HISTORY OF HUMANITY (1999).
[24] *A Trial of Witches,* 6 Howell State Trials 687, 697 (1665).

This expert testimony "concerning the bewitching of the children" was presented at one of the infamous Bury St. Edmunds witch trials in Suffolk, England, in 1665. The defendants were Rose Cullender and Amy Duny, two elderly widows accused of witchcraft by their neighbors. Their alleged crimes included the "bewitching" of several young village children. Dr. Brown purportedly used his expertise to offer the jurors an explanation, based on his "great knowledge," for the "strange fits" afflicting children in the village. Consequently, the jury convicted Rose Cullender and Amy Duny of witchcraft, and they were both hanged days later.[25] This episode illustrates a recurring problem in the law of evidence: courts have long struggled to distinguish between genuinely reliable expert knowledge and popularly persuasive but methodologically flawed "scientific" claims.

The practice of offering testimony in the form of an expert opinion became gradually entrenched in the common law, in both civil and criminal proceedings. Many legal historians cite the 1782 case *Folkes v. Chadd* as one of the earliest instances where there was a formal judicial recognition in support of expert opinion testimony.[26] This case involved the structural erosion of a harbor, and an expert was called to testify as to whether this decay was caused by a newly built sea bank. The King's Bench found that this question necessitated testimony from a "Mr. Smeaton," who apparently was an engineer possessing specialized knowledge regarding "the construction of harbours, the causes of their destruction, and how remedied."[27] His testimony was objected to on the grounds that he would "speak, not as to facts, but as to opinion," a precursor to Rule 701. Lord Mansfield ruled the opinion was admissible, regardless:

---

[25] *Id.* at 701–702.
[26] Folkes v. Chadd, 99 Eng. Rep. 589, 3 Dougl. K.B. 157 (1782); s*ee, e.g.,* Christopher M. Milroy, *A Brief History of the Expert Witness*, 7 ACAD. FORENSIC PATHOLOGY INT'L 516, 518 (2017).
[27] Folkes v. Chadd, 99 Eng. Rep. 589, 590, 3 Dougl. K.B. 157, 158–160 (1782).

> That opinion, however, is deduced from facts which are not disputed – the situation of banks, the course of tides and of winds, and the shifting of sands. . . . In matters of science no other witnesses can be called. The question then depends on the evidence of those who understand such matters. . . . The cause of the decay of the harbour is also a matter of science, and still more so, whether the removal of the bank can be beneficial. Of this, such men as Mr. Smeaton alone can judge. Therefore we are of opinion that his judgment, formed on facts was very proper evidence.[28]

Lord Mansfield also noted that "the whole case is a question of opinion, from facts agreed upon."[29] There was no dispute that decay had occurred at the harbor and that the plaintiff was injured. The only question for the court was one of causation: what caused the decay? Once this matter was resolved, the court's judgment would necessarily follow. This dispositive significance of expert opinion testimony remains true in many cases of modern litigation, as explained by one leading critic of "junk science" in the courtroom:

> In scientifically complex cases, experts are often, for all practical purposes, the entire case. If the plaintiffs' proffered expert testimony is excluded there will be no trial at all; the case will be decided summarily for the defense.[30]

Clearly, the stakes are high. And prior to the adoption of the FRE in 1975, courts were largely on their own to wade through the common law and figure out the framework for these high-stakes determinations. This led some commentators to describe the state of evidence law in the United States prior to 1975 as being a "crazy quilt" of disjointed and contradictory rulings.[31] Disputes were settled using little more than "dueling citations to outdated precedent or ambiguous statements in the common law."[32] The most significant such ruling was *Frye v.*

---

[28] *Id.*
[29] *Id.*
[30] Peter Huber, *Junk Science in the Courtroom,* 26 VAL. U. L. REV. 723, 727 (Summer 1992).
[31] David E. Bernstein & Eric G. Lasker, *Defending Daubert: It's Time to Amend Federal Rule of Evidence 702*, 51 WM. & MARY L. REV. 1, 47 (October 2015).
[32] Comm. On Rules of Practice and Procedure, Judicial Conference of the United States, *Rules of Evidence: A Preliminary Report on the Advisability and Feasibility of Developing Uniform Rules of Evidence for the United States District Courts*, 30 F.R.D. 73, 81 (1962).

*United States*, which gave courts the eponymous "*Frye* test" for admissibility for expert testimony.[33]

At the time the *Frye* decision was handed down in the early 20th century, courts were being confronted with new fields of science that were making their way into the courtroom via expert testimony. The growing presence of expert testimony was viewed as "anomalous" in the adversarial system of justice, and legal scholars engaged in debates on how to best absorb the evolving role of the expert witness.[34] One of the most significant expert testimony controversies that developed in American courts concerned the proposed introduction of results from a "systolic blood pressure test for deception," or what is commonly known today as a "lie detector" test.[35] The device purported to provide an objective assessment of whether an individual is stating a truth or a lie, based on contemporaneously "measuring and recording the blood pressure" of the declarant.[36] An early proponent of this technique was William Moulton Marston, a Professor of Psychology at American University. Marston, who also held a law degree from Harvard University and briefly worked as an attorney, had long been a critic of the common methods of discerning truth in the courtroom. In particular, he viewed testimony in the form of eyewitness recollections as intractably unreliable. For criminal cases, he considered a lie detector test administered to the defendant to be a superior means of uncovering the truth.[37]

---

[33] Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).
[34] *See, e.g.,* Learned Hand, *Historical and Practical Considerations Regarding Expert Testimony,* 15 HARV. L. REV. 40, 50–53 (1901) (describing the expert witness as an "anomaly" because such witness "is in effect not telling of facts at all, but of uniform physical rules, natural laws, or general principles, which the jury must apply to the facts").
[35] Lepore, *supra* note 16, at 1098; Victor S. Alpher & Richards Blanton, *The Accuracy of Lie Detection: Why Lie Tests Based on the Polygraph Should Not Be Admitted into Evidence Today,* 9 L. & PSYCH. REV. 67 (1985).
[36] Lepore, *supra* note 16, at 1138 (citing Brief for the Appellee at 5, in Frye v. United States, 293 F. 1013 (D.C. Cir. 1923)).
[37] *Id.* at 1096, 1115.

Utilizing his scientific expertise in psychology, Marston began researching and experimenting with new objective tests for deception, in particular, "the measurement of systolic blood pressure of a suspect while he is testifying."[38] Marston carried out testing of this method at a Harvard laboratory and reported that the test displayed "100% accuracy of judgment under very difficult conditions."[39] Marston also reported perfect accuracy of his lie detection test when administered to twenty "actual defendants in certain criminal courts."[40] However, in his publication's conclusion, Marston acknowledged that the "legal application of these tests presents a very interesting problem."[41]

Marston was determined to have this "interesting problem" resolved. He believed he found the perfect test case when he learned of the trial of James Alphonso Frye. On March 10, 1922, Frye was indicted for first-degree murder for the 1920 killing of Robert Wade Brown, a prominent doctor who was, by some accounts, one of the richest men in Washington D.C.[42] Although Frye confessed to the murder following an unrelated arrest for robbery, Frye now insisted that the confession was false.[43] Frye's fate largely hinged on his ability to convince a jury that he was now being truthful at trial in disclaiming his prior confession. Marston visited Frye in jail and questioned him about the confession while Frye was connected to a blood pressure lie detector test. Frye's defense attorneys then wanted to "offer the blood-pressure record which was made contemporaneously with the examination at the jail" and also "put [Marston] on the stand for the purpose of testifying as to the deception or nondeception of [Frye]

[38] William M. Marston, *Psychological Possibilities in the Deception Tests*, 11 J. OF CRIM. L. AND CRIMINOLOGY 551, 552 (1921).
[39] *Id.* at 553.
[40] *Id.*
[41] *Id.* at 570.
[42] Lepore, *supra* note 16, at 1119.
[43] *Id.* at 1124.

during the examination." Presiding Chief Justice Walter I. McCoy rejected the evidence, ruling that "the science has not sufficiently developed detection of deception by blood pressure to make it a usable instrument in a court of law."[44] Justice McCoy also ruled "it is for the jury to determine whether or not" Frye was being truthful.[45] Frye was subsequently convicted of murder.

Frye appealed his conviction to the Court of Appeals of District of Columbia, alleging a single error for the exclusion of this expert testimony. The appellate court upheld the conviction, and wrote what became known as the eponymous "*Frye* test":

> Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.[46]

Over the following decades, the *Frye's* "general acceptance in the field" test was adopted in "practically all of the courts of [the United States] which have considered the question of the admissibility of new scientific evidence."[47]

Although the *Frye* test became the dominant approach for determining admissibility of expert testimony, it was eventually replaced with a more "flexible" approach with the adoption of the FRE in 1975 and a series of landmark Supreme Court cases in the late 1990's.[48] At the time of enaction, the FRE included Rule 702, "Testimony by Experts." This Rule outlined the

---

[44] *Id.* at 1128, 1132.
[45] *Id.*
[46] Frye v. United States, 293 F. 1013, 1013 (D.C. Cir. 1923)
[47] State v. Washington, 622 P.2d 986, 991 (Kan. 1981).
[48] Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 580 (1993); Jim Hilbert, *The Disappointing History of Science in the Courtroom:* Frye, Daubert, *and the Ongoing Crisis of "Junk Science" in Criminal* Trials, 71 OKLA. L. REV. 759, 770 (2019).

admissibility of testimony based on "scientific, technical, or other specialized knowledge."[49] Because Rule 702 did not merely recite *Frye's* "general acceptance in the field" test, it created "sharp divisions" among the federal courts as to whether this new codified Rule operated in parallel with the *Frye* test, or supplanted it.[50] In 1992, the Supreme Court granted certiorari to settle this question in what became one of the most cited cases in federal court history, *Daubert v. Merrell Dow Pharmaceuticals, Inc.*[51]

*Daubert* involved plaintiffs who were suing the manufacturers of a drug known by the trade name Bendectin. The lawsuit alleged that while plaintiffs were *in utero,* their mothers took Bendectin to treat nausea, and that their prenatal exposure to the drug was the cause of their birth defects. The trial court dismissed the plaintiff's claim on summary judgment, ruling that "epidemiological studies are the most reliable evidence of causation in this area" and that "no epidemiological study ever performed has concluded that the use of Bendectin by pregnant women has a statistically significant association to birth defects."[52] Thus, the "animal-cell studies, live-animal studies, and chemical-structure analyses on which petitioners had relied" could not be introduced and create a fact issue.[53] The U.S. Court of Appeals for the Ninth Circuit affirmed, citing *Frye,* and finding that "animal and chemical studies . . . provide insufficient foundation," especially in light of "the massive weight" of the epidemiological studies that supported defendant's position.[54]

---

[49] Federal Rules of Evidence, Pub. L. No. 93-595, 88 Stat. 1926, 1937 (Jan. 2, 1975).
[50] *Daubert*, 509 U.S. at 585.
[51] *Id*. For the claim that *Daubert* became one of the most cited cases in federal court history, *see* Christopher J. Walker, *Most Cited Supreme Court Administrative Law Decisions,* YALE J. ON REGUL. (Oct. 9, 2014), https://www.yalejreg.com/nc/most-cited-supreme-court-administrative-law-decisions-by-chris-walker/.
[52] Daubert v. Merrell Dow Pharmaceuticals, Inc., 727 F. Supp. 570, 575 (S.D. Cal. 1989).
[53] *Daubert,* 509 U.S. at 584.
[54] Daubert v. Merrell Dow Pharmaceuticals, Inc., 951 F.2d 1128, 1130, 1131 (9th Cir. 1991).

But on review, the Supreme Court held that the lower courts had hemmed too tightly to the *Frye* approach. The FRE had "displaced" the single-pronged *Frye* test. Going forward, the *Frye* test "should not be applied in federal trials," and instead, federal courts should be guided by the FRE, with the "primary locus" of rules concerning expert testimony being found in Rule 702.[55] Notwithstanding the Court's admonition at the continued reliance on *Frye,* the Court provided guidance for lower courts for applying Rule 702, and this guidance still included *Frye*'s "widespread acceptance" test, but now it was but one factor in the analysis:

> Many considerations will bear on the inquiry [of whether the testimony's underlying reasoning or methodology is scientifically valid], including whether the theory or technique in question can be (and has been tested), whether it has been subjected to peer review and publication, its known or potential error rate and the existence and maintenance of standards controlling its operation, and whether it has attracted widespread acceptance within a relevant scientific community." The inquiry is a flexible one, and its focus must be solely on principles and methodology, not on the conclusions they generate. Throughout the judge should also be mindful of other applicable Rules.[56]

The Supreme Court held that Rule 702 now afforded parties greater flexibility than *Frye's* single criterion, and the matter was remanded to the lower courts to conduct a new inquiry that relied on Rule 702, as interpreted by the Court in *Daubert.* As a result of the *Daubert* decision, these inquiries to decide admissibility of expert testimony often take the form of complex evidentiary hearings in the trial court. These proceedings became such an entrenched, routine, and important procedure in litigation that they are now eponymously described as "*Daubert* hearings."[57]

---

[55] *Daubert,* 509 U.S. at 589.
[56] *Daubert,* 509 U.S. at 580.
[57] Munia Jabbar, *Overcoming* Daubert*'s Shortcomings in Criminal Trials: Making the Error Rate the Primary Factor in* Daubert*'s Validity Inquiry*, 85 N.Y.U. L. REV. 2034, 2042 (2010).

The Supreme Court would soon further develop Rule 702 in two other significant cases. In *General Electric Company. v. Joiner,* the Court held that rulings regarding the admissibility of expert scientific evidence should be reviewed on appeal using an abuse of discretion standard.[58] The *Joiner* court further held that it was not an abuse of discretion for the trial court to reject an experts' reliance on animal model evidence that was "so dissimilar to the facts presented" in the case at hand.[59] And then in *Kumho Tire Co. v. Carmichael,* the court ruled that *Daubert's* framework should be applied for all varieties of expert testimony, including those experts "who are not scientists."[60] *Daubert, Joiner, and Kumho* became foundational cases known as the "*Daubert* trilogy." The Court summarized this trilogy as providing courts and litigants with "exacting standards of reliability" with respect to admissibility of expert testimony.[61] Nevertheless, trial courts continued to display inconsistencies with how stringent or liberal they were applying Rule 702's requirements. This persistent confusion led to significant amendments to the language of Rule 702.[62]

In 2000, to address lingering confusion among trial courts and litigants, the essential holdings from the *Daubert* trilogy became codified into the amended text of Rule 702.[63] But even after this amendment, federal courts continued to implement "significantly more lenient standards for expert testimony than Rule 702 permits."[64] Many federal circuits were adjudicating disputes about the reliability of a particular methodology by concluding that such disputes speak

---

[58] Gen. Elec. Co. v. Joiner, 522 U.S. 136, 136 (1997).
[59] *Id.*
[60] Kumho Tire Co. v. Carmichael, 526 U.S. 137, 138 (1999).
[61] Weisgram v. Marley Co., 528 U.S. 440, 455 (2000).
[62] David E. Bernstein & Eric G. Lasker, *Defending Daubert: It's Time to Amend Federal Rule of Evidence 702*, 51 WM. & MARY L. REV. 1, 6 (October 2015).
[63] David E. Bernstein, *The Misbegotten Judicial Resistance to the* Daubert *Revolution,* 89 NOTRE DAME L. REV. 27, 29 (2013); FED. R. EVID. 702 advisory committee's note to 2000 amendment.
[64] David E. Bernstein, *The Misbegotten Judicial Resistance to the* Daubert *Revolution,* 89 NOTRE DAME L. REV. 27, 30 (2013); FED. R. EVID. 702 advisory committee's note to 2000 amendment.

to the *weight* of evidence rather than *admissibility*.[65] Some courts were taking such a "flexible" approach to *Daubert* that it amounted to an "abdication of their gatekeeper role."[66] This led to a wide range of approaches and standards for admissibility of expert testimony, which resulted in "roulette wheel randomness" in the federal courts as to when evidence would be deemed admissible.[67] This thwarted one policy goal of Rule 702, which is to ensure that "junk science" is not introduced in court.[68]

In 2023, the FRE were amended to again underscore the importance of federal courts acting in their "active gatekeeper" role under Rule 702. The default standard for deciding admissibility under all of FRE's requirements has long been by a preponderance of the evidence.[69] But to address the problem of some courts taking too permissive of an approach to Rule 702, leading to "roulette wheel" randomness, the amended text now expressly reaffirms that the burden rests with "the proponent [to demonstrate] to the court that it is more likely than not" that Rule 702's requirements have been satisfied.[70] Especially after this 2023 amendment, trial courts are on notice that they must take the time to carefully scrutinize any proffered expert testimony to ensure that the proponent has made a showing that it complies with Rule 702's requirements. The scope of this scrutiny must necessarily include the techniques and theories that underpin the testimony.

---

[65] Archibald Cruz, *The Paradigm Shift in the Proposed Amendment to Federal Rule of Evidence 702,* 75 BAYLOR L. REV. 265, 266 (2023) (citing as examples United States v. Shea, 211 F. 3d 658, 668 (1st Cir. 2000) and Puga v. RCX Sols., Inc., 922 F. 3d 285, 294 (5th Cir. 2019)).
[66] Victor E. Schwartz & Cary Silverman, The Draining of Daubert and the Recidivism of Junk Science in Federal and State Courts, 35 HOFSTRA L. REV. 217, 218 (Fall 2006).
[67] *Id.*
[68] *See generally* PETER W. HUBER, GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM (1991).
[69] *E.g.,* Bourjaily v. United States, 483 U.S. 171, 171 (1987) ("The traditional requirement [is that] that such [evidentiary] questions be established by a preponderance of proof").
[70] FED. R. EVID. 702.

Because Rule 702 requires that expert testimony be grounded in reliable methodology, the admissibility of animal model evidence ultimately depends on whether the animal modeling paradigm itself satisfies this reliability requirement. Thus, when expert testimony based on animal model evidence is being proffered, courts must scrutinize the animal model paradigm to ensure that it is a "reliable method." Based on what is known and understood about the lack of predictive value of animal models, proponents will typically be unable to make this showing, and animal model evidence should be considered by trial courts to be presumptively inadmissible.

## III. The Scientific Failure of Animal Models

The use of animals is a ubiquitous practice at biomedical and life sciences research institutions, but it is difficult to ascertain precise statistics. This is because the only federal law which regulates the use of animals in research, the Animal Welfare Act (AWA), statutorily does not define rats, mice, or birds as "animals."[71] These are the most common species used in research.[72] One study extrapolated statistics for other species, which are counted, to form an estimate of the total number of animals used in research, including mice, rats, and birds. This estimate concluded that there are approximately 111.5 million animals used in laboratory research in the United States each year, the overwhelming majority of which are mice and rats.[73]

Animals are used in a wide range of laboratory and scientific contexts, including as pedagogical tools in education, as a resource for producing biological products (e.g., insulin, heart valve replacements derived from pigs, or monoclonal antibodies), and as "models" for

[71] 7 U.S.C. § 2132(g).
[72] GENEVIEVE K. CROFT, CONG. RSCH. SERV., R47179, THE ANIMAL WELFARE ACT: BACKGROUND AND SELECT ISSUES 3 (July 14, 2002).
[73] Larry Carbone, *Estimating Mouse and Rat Use in Laboratories by Extrapolation from Animal Welfare Act-Regulated Species*, 11 SCI. REPORTS 493 (2021).

human biology.[74] "Animal modeling," sometimes called "comparative medicine," is a practice and paradigm that rests on an assumption that because "other animal species share physiological, behavioral, or other characteristics with humans," nonhuman animals can be used in laboratory experiments and observational studies with the aim of better understanding human biology.[75] It is solely this last category of laboratory animal use, "animal modeling," that falls under the scope of this article's critique. And although the practice of animal experimentation has been the subject of significant ethical controversy, this article does not offer any moral judgments or normative claims about the practice of using nonhuman animals in research.[76] This article is concerned solely with the epistemological utility of animal models, and its implications for the law of evidence.

Animal modeling is a practice that dates back thousands of years to ancient Greece. There, dissections of living animals were used to acquire basic biological knowledge, such as the principle that the "heart functions as a pump."[77] By the beginning of the twentieth century, "the use of animal modeling had increased dramatically," with the mouse being described as "the powerhouse for biomedical research."[78] The National Institutes of Health (NIH) is "the largest and most influential and complex biomedical research sponsor in the world."[79] According to a 2012 estimate from the Office of Extramural Research at the NIH, approximately 47% of all

---

[74] *Why Animals Are Used in* Research, NATIONAL INSTITUTES OF HEALTH (Aug. 16, 2024), https://grants.nih.gov/policy-and-compliance/policy-topics/air/why-animals-are-used-in-research.
[75] Aaron C. Ericsson, Marcus J. Crim, Craig L. Franklin, *A Brief History of Animal Modeling,* 110 MO. MED. 201, 201 (2013).
[76] For discussion of the ethical controversy, *see, e.g.*, PETER SINGER, ANIMAL LIBERATION NOW: THE DEFINITIVE CLASSIC RENEWED (2023).
[77] Ericsson, Crim, & Franklin, *supra* note 75, at 202.
[78] *Id.* at 201–203.
[79] *Review of National Institutes of Health Structure, Policies, and Programs, in* A NEW VISION FOR WOMEN'S HEALTH RESEARCH 85 (Sheila P. Burke, Alina Salganicoff, & Amy Geller eds., 2025).

grants now "have an animal research-based component."[80] Underpinning this dramatic expansion was the "assumption" that "animal research will be of benefit to humans."[81] Now, the use of animal models is often described by the largest and most well-known research enterprises in the United States as being "essential to major advances in medicine,"[82] "crucial for the acquisition of basic knowledge in biology,"[83] and "critical" for "designing and testing new therapies."[84] Such sweeping claims in defense of animal research have even become "axiomatic" to some researchers. [85]

But the evidence for these lofty pronouncements is lacking. To be sure, animal models have played a role in certain historical scientific advances, and they remain widely used in biomedical research. However, their predictive limitations—particularly in complex human disease and drug response—are increasingly recognized. Writing in *BMJ* in 2004*,* a team of researchers conducted a meta-analysis of published systematic reviews of animal studies.[86] The goal of such systematic reviews is to "compare the results of [animal studies] with the results of the corresponding clinical trials." But the researchers found that such systematic reviews were rarely occurring at all. And the researchers found that those "few existing reviews [that did occur] have highlighted deficiencies" with the use of animal models in biomedical research.[87]

---

[80] INSTITUTE OF MEDICINE AND NATIONAL RESEARCH COUNCIL, INTERNATIONAL ANIMAL RESEARCH REGULATIONS 23, (2012).
[81] Aysha Karim Kiani et al., *Ethical Considerations Regarding Animal Experimentation,* 63 J. PREVENTIVE MED. & HYG E255, E256 (2022).
[82] *Animal Care and Use at Johns Hopkins Medicine,* Johns Hopkins University, https://www.hopkinsmedicine.org/research/resources/offices-and-policies/animal-care (last visited Apr. 21, 2026).
[83] NATIONAL RESEARCH COUNCIL, USE OF LABORATORY ANIMALS IN BIOMEDICAL AND BEHAVIORAL RESEARCH 27 (1988).
[84] *Animal Research at HMS*, Harvard Medical School, https://hms.harvard.edu/research/animal-research (last visited Apr. 21, 2026).
[85] Pandora Pound et al., *Where is the Evidence That Animal Research Benefits Humans?*, 328 BMJ 514, 514 (Feb. 28, 2024).
[86] *Id.*
[87] *Id.* at 517.

Similarly, in 2008, the *Journal of the Royal Society of Medicine* published an analysis of the claim that animal experimentation was "indispensable" to medical advancement. It found that the evidence to support this claim was "lacking in logical or evidential support."[88] The author aptly observed that the implicit purpose of animal models is to have "predictive value" for results in humans.[89] Researchers create disease conditions or test an investigational drug in a laboratory animal because they ostensibly have inferential value. The implicit logic underpinning animal models is that we can reasonably expect the results that are observed in laboratory animals will usually correspond to similar conditions or exposures in humans. But this predictive power is often merely assumed as an article of faith.

There have been "relatively few" actual quantitative studies aimed at validating the predictive value of animal models.[90] Those studies that have been performed suggest that the evidential weight of the animal models was "at best inconclusive, and sometimes wholly misleading."[91] Taken together, those studies that have been conducted "provide no statistically credible evidence that these animal models contribute any predictive value, either separately or in combination."[92] Strong evidence suggests that use of animal models "has not worked in any predictable, reliable, reproducible, or translatable sense to advance human medicine and conquer human diseases."[93] These scientific limitations are not merely academic. They have direct implications for courts applying Rule 702, which requires that expert testimony be based on

---

[88] Robert A.J. Matthews, *Medical Progress Depends on Animal Models — Doesn't It?*, 101 J. ROYAL SCI. MED. 95, 97 (2008).
[89] *Id.* at 96.
[90] *Id.*
[91] *Id.* at 97.
[92] *Id.*
[93] John J. Pippin, *Animal Research in Medical Sciences: Seeking a Convergence of Science, Medicine, and Animal* Law, 54 S. TEX. L. REV. 469, 504 (2013).

reliable principles and methods.[94] Animal model evidence is too unreliable to meet this requirement.

## IV. Regulatory Embrace and Retreat from Animal Models

Despite this lack of scientific validation or foundation, the attitude that animal research is necessary for medical advancements has been reflected in federal policy for nearly a century. In 1938, Congress passed the Federal Food, Drug, and Cosmetic Act (FD&C Act).[95] This law required that pharmaceutical companies submit data to regulators with the U.S. Food and Drug Administration (FDA) demonstrating a drug's safety prior to marketing.[96] This safety data often came in the form of tests on animals, including mice, rats, guinea pigs, dogs, and monkeys.[97] The new law came on the heels of a highly publicized disaster where a deadly formulation of the drug sulfanilamide was sold on the market.[98] However, this original 1938 version of the (FD&C Act) contained an exemption for drugs administered to people in clinical trials.[99]

In 1962, Congress expanded the scope of the FD&C Act with an amendment sponsored by Senator Estes Kefauver of Tennessee. The "Kefauver amendment" came in response to another widely publicized pharmaceutical disaster, involving the drug thalidomide. Thalidomide was a hypnotic used to treat a range of conditions, including nausea associated with pregnancy, commonly known as "morning sickness." Although thalidomide was never approved for use in

---

[94] FED. R. EVID. 702(c).
[95] Federal Food, Drug, and Cosmetics Act, Pub. L. No. 75-717, 52 Stat. 1040 (1938).
[96] Sharon J. Jacobs, *Crises, Congress, and Cognitive Biases: A Critical Examination of Food and Drug Legislation in the United* States, 64 FOOD & DRUG L.J. 599, 607 (2009).
[97] *Id*.
[98] *Drug Fatality Cause is Traced to Elixir; A.M.A. Chemists Say Diethylene Glycol Added to Sulfanilamide Killed 13,* N.Y. TIMES, Oct. 20, 1937.
[99] Federal Food, Drug, and Cosmetic Act, Pub. L. No. 75-717, 52 Stat. 1040, 1053, § 505(i) (1938) ("The Secretary shall promulgate regulations for exempting from the operation of this section drugs intended solely for investigational use by experts qualified by scientific training and experience to investigate the safety of drugs.").

the United States, the drug caused 5,000 to 6,000 birth defects in Germany, where the drug was widely used.[100] Despite the fact that it never received FDA approval, thalidomide was distributed for "experimental or investigational use" to about 1,200 physicians in the United States.[101] After the passage of the Kefauver amendment, pharmaceutical companies in the U.S. were required to seek FDA approval prior to any clinical trials of an investigational drug. In order to get this approval, data had to be submitted to the FDA in the form of "preclinical tests (including tests on animals)" to "justify the proposed clinical testing."[102] After this "preclinical" testing in animals, drug companies were required to test a drug's safety and efficacy using a three-step series of "phased" clinical trials, with each phase using a wider and larger pool of (human) volunteer participants. The underlying rationale to this arrangement seems obvious. If problems including toxicity or teratogenicity (propensity for causing birth defects) could be detected in a new drug using mice or other laboratory animal testing, the FDA could withhold that drug from clinical trials and the market to prevent harmful outcomes for humans, like those that occurred with sulfanilamide and thalidomide.

Although this rationale may seem intuitive, the evidence reveals that it is fundamentally flawed. The safety and efficacy profile of a drug is the result of a constellation of factors, including the ways in which that drug is absorbed, distributed, metabolized, and excreted in the body. The careful analysis of these factors is known as pharmacokinetics, an integral part of the drug discovery process.[103] Pharmacokinetic processes are all mediated and determined at the

---

[100] Howard A. Rusk, *Thalidomide Tragedy — Surge in Congenital Defects Poses Questions on Precautions and Care*, N.Y. TIMES, Aug. 5, 1962.
[101] *U.S. Considering Regulations to Reduce Peril of New Drugs,* N.Y. TIMES, July 29, 1962.
[102] Kefauver amendment to the Federal Food, Drug, and Cosmetic Act, Pub. L. No. 87-781, 76 Stat. 783; Sharon J. Jacobs, *Crises, Congress, and Cognitive Biases: A Critical Examination of Food and Drug Legislation in the United* States, 64 FOOD & DRUG L.J. 599, 611 (2009).
[103] Yuhua Li et al., *Current Trends in Drug Metabolism and Pharmacokinetics,* 9 ACTA PHARMACEUTICA SINICA B 1113, 1114 (2019).

genetic and molecular level, which is precisely the level at which a mouse becomes a mouse and a human becomes a human. As a result, drugs frequently do not exhibit the same pharmacokinetics in animals as they do in humans, a problem described as "poor translatability."[104] This means that drugs do not have the same safety and efficacy profile across species, undermining the predictive value of animal models. This creates the potential for two types of errors in the drug development process: false negatives and false positives.

"False negatives" are drugs that appear safe and effective in animals but go on to have no efficacy or cause harm in people. That harm can become apparent either in clinical trials or when it is released on the market, leading to a drug being relabeled or recalled. For example, the experimental drug TGN1412 was in development to treat a wide range of autoimmune diseases, including multiple sclerosis, rheumatoid arthritis, and certain cancers.[105] During preclinical testing on monkeys, TGN1412 was demonstrated to be "safe and efficacious."[106] But when six human volunteers received the drug as part of a clinical trial in London, within 90 minutes they experienced "a systemic inflammatory response" and became critically ill.[107] These adverse clinical outcomes occurred in these volunteers despite the fact that the monkeys safely received a dose 500 times greater than that given to the volunteers.[108] Even after researchers knew what they were looking for, "they could not duplicate these responses" in monkeys, "suggesting that TGN1412 does not behave the same in nonhuman primate species as in humans."[109]

---

[104] Lindsay J. Marshall et al., *Poor Translatability of Biomedical Research Using Animals – A Narrative Review*, 51 ALTERNATIVES TO LAB'Y ANIMALS 102, 102 (2023).
[105] Sarah Bosley, Jacqueline Maley, and Suzanne Goldenberg, *Doctors Call for Worldwide Help in Race to Save Lives of Men Who Collapsed 'Like Dominoes' in Drug Trial*, THE GUARDIAN (Mar. 16, 2006), https://www.theguardian.com/science/2006/mar/17/frontpagenews.medicineandhealth
[106] Attarwala H, *TGN142: From Discovery to Disaster*, 2 J. YOUNG PHARM. 332, 333 (2010).
[107] E. William St. Clair, *The Calm After the Cytokine Storm: Lessons From the TGN1412 Trial,* 118 J. CLINICAL INVESTIGATION 1344, 1345 (2008).
[108] *Id*.
[109] *Id.* at 1346.

False negatives are a common occurrence. The FDA has reported that 92% of drugs entering clinical trials ultimately fail or cause harm in humans, never to be released on the general consumer market.[110] Every one of these drugs made it through preclinical testing where they appeared safe and effective in animals. Even in the case of thalidomide, the Kefauver amendment's required animal testing would not have protected the public. *Post hoc* tests carried out on laboratory animals who were administered thalidomide have repeatedly failed to reproduce the birth defects that are observed in humans.[111] Researchers could not even reproduce the harmful birth defects in rats, which was considered "more disturbing when one realizes that the pregnant rat is a laboratory animal commonly used to screen drugs for teratogenic effects."[112]

"False positives" are drugs that appear unsafe or ineffective in animals but are actually safe and clinically beneficial in people. Perhaps the most well-known example of a "false positive" is the world's first antibiotic, penicillin. In 1929, Alexander Fleming was experimenting with penicillin when he saw that it was effective at killing harmful bacteria in a petri dish.[113] He then tested it on rabbits, where it was ineffective. We now understand that this is because rabbits excrete penicillin in their urine before the drug can be pharmacologically active. As a result of these failed tests on rabbits, penicillin was shelved for over a decade. Later, the drug was tried on a human patient for whom there was no other therapy. It worked.[114] If penicillin had been first tested on guinea pigs, it may have never been tested on humans, as the

---

[110] U.S. FOOD AND DRUG ADMINISTRATION, INNOVATION/STAGNATION: CHALLENGE AND OPPORTUNITY ON THE CRITICAL PATH TO NEW MEDICAL PRODUCTS 8 (2004).
[111] Robert L. Brent, *Drug Testing in Animals for Teratogenic Effects: Thalidomide in the Pregnant Rat*, 64 J. PEDIATRICS 762 (1964).
[112] *Id*. at 763.
[113] C. RAY GREEK & JEAN SHINGLE GREEK, SACRED COWS AND GOLDEN GEESE: THE HUMAN COST OF EXPERIMENTS ON ANIMALS 73–74 (2000).
[114] *Id*.

drug is deadly to these animals, even at doses "well below the toxic level for other animals."[115] Fleming later stated about the history of the development of penicillin that it was "fortunate we didn't have these animal tests in the 1940s, for penicillin would probably never [have] been granted a license, and possibly the whole field of antibiotics might never have been realised."[116] It is hard to quantify these "false positives," because some effective therapies may have never been tried in humans, having been abandoned during the development process due to misleading tests in animals.[117]

The U.S. Department of Health and Human Services (HHS) oversees most of the federal government's involvement in the domains of biomedical research, pharmaceutical development, and public health. HHS agencies include the Centers for Disease Control and Prevention (CDC), the FDA, and the NIH. In 1994, HHS released an official statement where it claimed that "[v]irtually every medical achievement of the last century has depended directly or indirectly on research with animals" and that "animal research is necessary to achieve medical progress."[118] But more than thirty years later, there are clear signs that this absolutist outlook is no longer shared by HHS and its constituent agencies. The federal government is showing increasing skepticism towards the scientific utility of animal models. That skepticism is reflected in seismic changes in federal policies and programs.

---

[115] Dorothy M. Hambre et al., *The Toxicity of Penicillin as Prepared for Clinical Use*, 206 AM. J. MED. SCI. 642, 642 (1943).
[116] Dennis V. Parke, *Clinical Pharmacokinetics in Drug Safety Evaluation*, ALTERNATIVES TO LAB'Y ANIMALS 207, 208 (1994).
[117] Lucy Meigs et al., *Animal Testing and Its Alternatives — The Most Important Omics is* Economics, 35 ALTEX 275, 298–88 (2018).
[118] U.S. Public Health Service of the Department of Health and Human Services, *The Importance of Animals in Biomedical and Behavioral Research*, 37 PHYSIOLOGIST 107 (1994).

Federal policy first changed with respect to the use of chimpanzees. Chimpanzees are the closest living relatives to humans, sharing approximately 98.8% of our genetic material.[119] This close similarity with humans has been cited as the basis for their use in scientific research, as well as the reason for significant ethical and public concern.[120] In 1986, the NIH launched the Chimpanzee Breeding and Research Program. Consequently, the use of chimpanzees in experiments, particularly those investigating HIV/AIDS, sharply increased, with 1,500 chimpanzees being used at six U.S. research institutions.[121] Almost immediately, there was public outcry against the practice, led by well-known public figures, including Dr. Jane Goodall.[122] This opposition was based primarily on ethical concerns, but researchers also criticized the use of chimpanzees in HIV/AIDS research due to its track record of "failure."[123] These critiques led to multiple bills being introduced in Congress that would federally prohibit the use of chimpanzees in research.[124]

In the face of this growing public and political pressure, Congress and the NIH asked the Institute of Medicine (IOM) to conduct a study and release a report regarding the utility of using chimpanzees in research.[125] In December 2011, the IOM released a landmark report, which concluded that "most current use of chimpanzees for biomedical research is unnecessary."[126]

---

[119] *DNA: Comparing Humans and Chimps,* AM. MUSEUM NAT. HIST., https://www.amnh.org/exhibitions/permanent/human-origins/understanding-our-past/dna-comparing-humans-and-chimps (last visited May 1, 2026).
[120] Jane Goodall, *A Plea for the Chimps*, THE NEW YORK TIMES MAGAZINE, May 17, 1987, at 108.
[121] NATIONAL RESEARCH COUNCIL, CHIMPANZEES IN RESEARCH: STRATEGIES FOR THEIR ETHICAL CARE, MANAGEMENT, AND USE 7 (1997).
[122] Goodall, *supra* note 120.
[123] Jarrod Bailey, *An Assessment of the Role of Chimpanzees in AIDS Vaccine Research,* 36 ALTERNATIVES TO LAB'Y ANIMALS 381, 419 (2008) ("[V]accines of many types have been tested in chimpanzees prior to clinical trials, and their correlation to, and predictive nature for, the human response is demonstrably poor.").
[124] Great Ape Protection Act, H.R. 5852, 110th Cong. (2008); Great Ape Protection Act of 2010, S. 3694, 111th Cong. (2010); Great Ape Protection and Cost Savings Act of 2011, H.R. 1513, 112th Cong. (2011).
[125] INSTITUTE OF MEDICINE AND NATIONAL RESEARCH COUNCIL OF THE NATIONAL ACADEMIES, CHIMPANZEES IN BIOMEDICAL AND BEHAVIORAL RESEARCH: ASSESSING THE NECESSITY 1 (2011).
[126] *Id.* at 4.

Although the report did not directly address the failure of chimpanzee models to provide reliable data, its release occurred around the same time that other researchers were publishing review articles questioning the reliability of chimpanzees as models of human disease.[127] Within an hour of the IOM report's release, Dr. Francis Collins, Director of the NIH, announced that the NIH would "not entertain grant applications involving chimpanzees until further notice."[128] Funding never resumed. In 2015, the NIH announced that it was making permanent its prohibition on funding chimpanzee research, and that it would be retiring the remaining chimpanzees in its custody to sanctuaries.[129]

The federal government continues to accelerate its movement away from animal models. In 2026, U.S. Department of Health and Human Services Secretary Robert F. Kennedy Jr. stated that the agency has a "commitment to replace animal testing with human-relevant, scientifically rigorous methods."[130] This announcement came on the heels of a CDC announcement that it was ending its use of monkeys in research, affecting experiments using approximately 200 macaque monkeys.[131] At the same time, the FDA also announced a new policy that "recommended using newer non-animal testing tools – known as new approach methodologies, or NAMs – when

---

[127] Andrew Knight, *The Poor Contribution of Chimpanzee Experiments to Biomedical* Progress, 10 J. APPLIED ANIMAL WELFARE SCI. 281 (2007); Jarrod Bailey, *An Assessment of the Role of Chimpanzees in AIDS Vaccine Research,* 36 ALTERNATIVES TO LAB'Y ANIMALS 381 (2008); Jarrod Bailey, *An Assessment of the Use of Chimpanzees in Hepatitis C Research Past, Present, and Future: 1 Validity of the Chimpanzee Model,* 38 ALTERNATIVES TO LAB'Y ANIMALS 387 (2010); Jarrod Bailey, *Lessons from Chimpanzee-based Research on Human Disease: The Implications of Genetic Differences*, 39 ALTERNATIVES TO LAB'Y ANIMALS 527 (2011); Ray Greek, Lawrence A. Hansen, & Andre Menache, An Analysis of the Bateson Review of Research Using Nonhuman Primates, 1 MEDICOLEGAL AND BIOETHICS 3 (2011).
[128] Meredith Wadman, *Most Biomedical Research on Chimpanzees Deemed 'Unnecessary,'* NATURE (Dec. 16, 2011) https://www.nature.com/articles/nature.2011.9663.
[129] Sara Reardon, *NIH to Retire All Research Chimpanzees*, NATURE (Nov. 18, 2015) https://www.nature.com/articles/nature.2015.18817.
[130] *U.S. F.D.A. Pushes for New Methods to Replace Animal Testing in Early Drug Studies*, REUTERS (Mar. 18, 2026 at 8:23 AM PDT) https://www.reuters.com/business/healthcare-pharmaceuticals/us-fda-pushes-new-methods-replace-animal-testing-early-drug-studies-2026-03-18/
[131] David Grimm, *Political Appointee Tells CDC to End Monkey Research*, 390 SCIENCE 868, (Nov. 27, 2023).

companies submit nonclinical safety data to support a drug application."[132] FDA Commissioner Marty Makary echoed Secretary of Kennedy's sentiment, saying that "[m]odern science has given us far more effective ways of evaluating drug safety than animal testing."[133] Also in 2026, the NIH announced that it was investing $150 million in human-based research. The agency stated that the goal of this new investment was "to reduce use of animal models" and "produce more predictive models of human disease."[134]

The movement away from animal models isn't only being spearheaded by the executive branch. In December 2022, Congress passed the FDA Modernization Act 2.0. This law allocated $5 million to the FDA to accelerate development of alternatives to animal testing. It also amended the FD&C to repeal the Kefauver amendment's requirement for animal testing.[135] The change represented "a new chapter in drug development," based in part on a recognition that "many medications that appear promising in animals do not pan out in humans."[136] Don Ingber, a Harvard University bioengineer was even more blunt about the reasons for the change: "[a]nimal models are wrong more often than they are right."[137]

## V. Animal Model Evidence in the Courtroom

There is a wide range of litigation where scientific questions related to human health and disease are central to factual issues that the jury is tasked with deciding. These include cases

---

[132] *U.S. F.D.A. Pushes for New Methods to Replace Animal Testing in Early Drug Studies*, REUTERS (Mar. 18, 2026 at 8:23 AM PDT) https://www.reuters.com/business/healthcare-pharmaceuticals/us-fda-pushes-new-methods-replace-animal-testing-early-drug-studies-2026-03-18/
[133] *Id.*
[134] National Institutes of Health, *NIH Invests $150 million in Human-Based Research to Reduce Use of Animal Models,* NIH.GOV (Mar. 18, 2026), https://www.nih.gov/news-events/news-releases/nih-invests-150-million-human-based-research-reduce-use-animal-models.
[135] FDA Modernization Act 2.0, Pub. L. No. 117-328, 136 Stat. 5821, § 3209, (2022).
[136] Emily Anthes, *Could the Next Blockbuster Drug Be Lab-Rat Free?*, N.Y. TIMES (Mar. 7, 2023) https://www.nytimes.com/2023/03/07/health/drug-animals-testing.html.
[137] Meredith Wadman, *FDA No Longer Needs to Require Animal Tests Before Human Drug Trials,* SCIENCE (Jan. 10, 2023), https://www.science.org/content/article/fda-no-longer-needs-require-animal-tests-human-drug-trials.

involving medical malpractice, product liability, and "toxic torts." When parties in litigation seek to introduce animal model evidence to support expert conclusions, courts will have to apply the requirements of Rule 702. Courts now have to conduct that analysis in light of a growing recognition that "animal models are wrong more often than they are right."[138] Such a methodology cannot be considered "reliable."

Despite this inherent unreliability, animal models continue to be used and defended to establish causation in tort litigation.[139] But the same unreliability of animal models which poses the dual problem of "false positives" and "false negatives" during drug development is transposed in the context of litigation, leading to unjust outcomes:

> There is no more important issue in the law of torts than factual causation. If a defendant is held liable for something it did not do, then the justice system has failed. On the other hand, if a defendant is able to convince a jury that it did not cause an injury for which it was responsible, that is also a miscarriage of justice. Both failures of our justice system are more likely to arise when factual causation is entwined with scientific evidence.[140]

As made clear by the adoption of the 2023 amendments to the FRE, courts have often been too permissive in their Rule 702 gatekeeping role.[141] In a permissive environment, each litigant will be permitted to offer their own expert testimony to explain the "science" that supports their side of the claim. This has been described as a "classic battle of the experts, a battle in which the jury must decide the victor."[142] The parties are incentivized to seek out persuasive and experienced "experts" who are willing, often for a fee, to present the "science"

---

[138] *Id.*
[139] *See, e.g.,* Kristen Ranges & Jessica Owley, *supra* note 8, at 306.
[140] Victor E. Schwartz & Cary Silverman, *The Draining of Daubert and the Recidivism of Junk Science in Federal and State Courts*, 35 HOFSTRA L. REV. 217, 217 (Fall 2006).
[141] *See generally* Mark A. Behrens & Andrew J. Trask, *Federal Rule of Evidence: A History and Guide to the 2023 Amendments Governing Expert Evidence*, 12 TEX. A&M L. REV. 43 (2024).
[142] Ferebee v. Chevron Chemical Co., 736 F.2d 1529, 1535 (1984) (citing Jenkins v. United States, 307 F.2d 637, 646 (D.C. Cir. 1962)).

that supports that party's case.[143] Methodologies and techniques that are pliable and easily manipulable will often be the trade of such experts-for-hire. Animal model evidence presents perhaps the most malleable form of "research" for manufacturing data that supports the party's position. Laboratory tests can be carried out on animals in controlled environments, in a near-infinite number of conditions and arrangements, utilizing dozens of different possible species. This presents an opportunity to experts-for-hire, who can selectively cherry pick from a wide range of results:

> If rats cope with the heaviest dose of a chemical that can be soaked into their food and water, you can always gavage them. Or try mice and rabbits. . . . [T]he only stopping rule is discovery of the sought-for effect, or exhaustion of the investigator (or his funds). Many of the risk assessment procedures used today are logically indistinguishable from those used by the Inquisition.[144]

When animal model evidence is admitted despite the fact that it is misleading and failing to predict the response in humans, in some cases that will be for the benefit of plaintiffs. In other cases, it will be for the benefit of defendants. But either way, the verdict and judgment that follows will be a mistake based on an error. And that mistake will have occurred because the court failed in its gatekeeping role. This failure undermines the purpose of the FRE, to "ascertain[] the truth and secur[e] a just determination:"[145]

The litigation surrounding Bendectin provides a particularly instructive case study of how animal model evidence can produce "false positives" and misleading testimony in the courtroom. Starting in 1980, Merrell Dow, the manufacturer of Bendectin, was the defendant in several toxic tort cases that went to trial. Early trials were spearheaded by legendary tort litigator Melvin

---

[143] *See generally,* PETER W. HUBER, GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM (1991).
[144] William C. Clark, *Witches, Floods, and Wonder Drugs: Historical Perspectives on Risk Management, in* SOCIETAL RISK ASSESSMENT 291 (1980).
[145] FED. R. EVID. 102, "Purpose."

Belli.[146] These cases alleged the drug was the cause of the plaintiffs' birth defects.[147] Some cases were tossed at summary judgment but others were allowed to proceed to trial despite the fact that the FDA had investigated the matter and concluded that "available data do not demonstrate an association between birth defects and Bendectin."[148] In response to the growing number of lawsuits and negative publicity, Merrell voluntarily withdrew Bendectin from the market in 1983.[149] Even after this voluntary withdrawal, the lawsuits continued. The success of each lawsuit was largely dependent on the evidentiary rulings on the expert testimony.[150]

Eventually, one of these rulings from a Bendectin trial, *Daubert v. Merrell Dow Pharmaceuticals*, made its way up to the Supreme Court. When analyzing the record from the trial court, the Supreme Court ruled that the lower courts were applying too rigid of a standard for determining admissibility of expert testimony. The lower courts should have given plaintiffs' proffered experts greater consideration. "Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence."[151] The Supreme Court seemed persuaded that the plaintiffs in *Daubert* had presented testimony from eight experts, "each of whom also possessed impressive credentials."[152] In the footnote to this statement, the Supreme Court singled out two particular experts. One was Shanna Helen Swan, who had "served as a consultant to the World

---

[146] M. Korcok, *The Bendectin Debate,* 123 CAN. MED. ASSOC. J. 922, 922 (1980).
[147] Joseph Sanders, *From Science to* Evidence: The Testimony on Causation in the Bendectin Cases, 46 STAN. L. REV. 1, 4 (1993).
[148] Daubert v. Merrell Dow Pharmaceuticals, Inc., 43 F.3d 1311, 1314 (9th Cir. 1995) (citing U.S. Dep't of Health & Hum. Svcs. News, No. P80-45 (Oct. 7, 1980)).
[149] Associated Press, *Company Stops Making Morning Sickness Drug,* N.Y. TIMES, June 10, 1983, at A16.
[150] Michael D. Green, *Legacies of the Bendectin Case*, 278 SCIENCE 1574, 1575 (1997); J. Sanders, *From Science to Evidence: The Testimony on Causation in the Bendectin Cases,* 46 STAN. L. REV. 1, 32 (1993).
[151] Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 596 (1993).
[152] *Daubert*, 509 U.S. at 583.

Health Organization, the Food and Drug Administration, and the National Institutes of Health."[153]

Swan was a frequent "expert for hire" for plaintiffs in toxic tort cases. At the time *Daubert* was decided, Swan had "taken the witness stand in about 50 cases against the makers of some of medicine's most notorious drugs and medical devices."[154] Swan's affidavit regarding the safety of Bendectin was part of the Joint Appendix in *Daubert.*[155] The affidavit stated her opinion that Bendectin was associated with birth defects, and the association was to a level of statistical significance. The basis for this opinion were animal studies and an "unpublished" meta-analysis of epidemiological studies. This 1993 affidavit stated that there was a publication "in progress" related to this epidemiological meta-analysis.[156] But based on this author's search of PubMed, the National Library of Medicine's database of biomedical scientific literature, it does not appear that any such meta-analysis ever did appear in a peer-reviewed publication.

Despite the Court's ruling in *Daubert* that the lower courts should have applied a more relaxed standard under Rule 702, on remand the circuit court once again ruled that the plaintiff's expert testimony was inadmissible.[157] Writing for the Ninth Circuit, Judge Kazinski observed that the plaintiff's experts failed to "point to some authority for extrapolating human causation from teratogenicity [causing birth defects] in animals."[158] *Daubert* didn't do anything to change this fatal flaw. Based on our current understanding of the safety of Bendectin, it appears the Ninth Circuit got it right (again). The overwhelming scientific consensus is now that Bendectin

---

[153] Daubert, 509 U.S. at 583 n. 2.
[154] Catherine L. Yap-Yang, *Under Attack: Testimony For Hire*, BLOOMBERG (Jan. 17, 1993), https://www.bloomberg.com/news/articles/1993-01-17/under-attack-testimony-for-hire
[155] Joint Appendix, Affidavit of Shanna Helen Swan, Ph.D., Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993) (Dkt. No. 187), 1992 WL 12012184, at *112aa.
[156] *Id.*
[157] Daubert v. Merrell Dow Pharmaceuticals, Inc., 43 F.3d 1311, 1319 (9th Cir. 1995).
[158] *Id.*

is not associated with any increased incidence of birth defects.[159] In 2013, the FDA approved a new drug called Diclegis, for "the treatment of nausea and vomiting of pregnancy in women." Diclegis has the same active ingredient as Bendectin. In its reporting about the drug's return to the market, *The Associated Press* now referred to the prior Bendectin tort cases—which had resulted in several multi-million dollar jury awards and the seminal *Daubert* case—as a mere "safety scare."[160] And it was "safety scare" predicated largely on a false positive based on unreliable animal model evidence.

Animal model evidence can also present erroneous "false negatives" that lead to case assessments and outcomes that incorrectly skew in favor of the defendant in toxic tort product liability cases. In 1950, a landmark study was published by Richard Doll and Austin Bradford Hill that established an association between cigarette smoking and lung cancer.[161] The study was a basic case-control design, that compared the tobacco consumption habits of more than 700 patients with lung cancer with an equal number of patients without lung cancer (controls). The study found that "the risk of developing [lung cancer] increases in simple proportion with the amount smoked, and that it may be approximately 50 times as great among those who smoke 25 or more cigarettes a day as among non-smokers."[162] This study was followed by another landmark study known as The British Doctors Study. This study followed 40,000 British physicians over decades as part of a prospective cohort study design, and concluded that smoking was strongly associated with lung cancer, heart disease, and reduced life expectancy.[163]

---

[159] Slaughter et al., *supra* note 11.
[160] *Morning Sickness Drug Returns,* N.Y. TIMES (Apr. 8, 2013), https://www.nytimes.com/2013/04/09/us/morning-sickness-drug-returns.html.
[161] Richard Doll & A. Bradford Hill, *Smoking and Carcinoma of the Lung*, 4682 BMJ 739 (1950).
[162] *Id.* at 747.
[163] Richard Doll & A. Bradford Hill, *The Mortality of Doctors in Relation to Their Smoking Habits,* 4877 BMJ 1451 (1954).

When this study was published in 1954, it summarized the state of the epidemiological science as already showing "that there are more heavy smokers and fewer non-smokers among patients with lung cancer than among patients with other diseases."[164]

Almost immediately, there were calls for labeling, regulation, and accountability for the tobacco industry for the harms caused by their products. The tobacco industry responded by using "sophisticated public relations approaches" to "erode, confuse, and condemn the very science that now threatened to destroy its prized, highly popular, and exclusive product."[165] Their strategy demonstrates how the limitations of scientific evidence, including variability in experimental and animal studies, can be exploited to manufacture doubt about causation.

The tobacco industry created the "Tobacco Industry Research Committee (TIRC)," with the goal of "securing" science that reflected favorably on their product.[166] The first scientific director of TIRC was Clarence Cook Little, who immediately went to work favoring "basic science investigations into the mechanism of cancer, often using animal models."[167] Researchers soon developed a "smoking inhalation machine" for the purpose of exposing animals in the laboratory to tobacco smoke and then studying the health effects.[168] Many of these studies failed to replicate the same link between tobacco and lung cancer that was observed in the human-based epidemiological studies.[169] As a result, even as late as the 1990s, tobacco companies were

---

[164] *Id.* at 1451.
[165] Allan M. Brandt, *Inventing Conflicts of Interest: A History of Tobacco Industry Tactics,* 102 AM. J. PUB. HEALTH 63 (2012).
[166] *Id.* at 66.
[167] *Id.* at 66–67.
[168] H. Baumgartner & C.R.E. Coggins, *Description of a Continuous-Smoking Inhalation Machine for Exposing Small Animals to Tobacco Smoke,* 10 CONTRIBUTIONS TO TOBACCO & NICOTINE RSCH. 169 (1980).
[169] Lisa Kramer & Ray Greek, Human Stakeholders and the Use of Animals in Drug Development, 123 BUS. & SOC'Y REV. 3, 11 (2018).

maintaining that tobacco was safe. In a 1993 deposition taken by attorney Stanley Rosenblatt, Philip Morris President and C.E.O. William Campbell addressed this issue:

> Mr. Rosenblatt: Does cigarette smoking cause cancer?
>
> Mr. Campbell: To my knowledge, it's not been proven that cigarette smoking causes cancer.
>
> Mr. Rosenblatt: What do you base that on?
>
> Mr. Campbell: I base that on the fact that traditionally, there is, you know, in scientific terms, there are hurdles related to causation, and at this time there is no evidence that – they have not been able to reproduce cancer in animals from cigarette smoking.[170]

Today, the link between cigarette smoking and lung cancer is well-established, and accepted by the United States Surgeon General, the World Health Organization, and other leading health authorities.[171] The failure to reproduce this phenomena in experiments on animals is an example of a "false negative." This failure demonstrates the inherent flaws of animal model evidence and proves why this type of evidence should be considered unreliable and inadmissible under Rule 702.

## VI. Toward a Rule of Presumptive Inadmissibility

Rule 702 and the *Daubert* trilogy require trial courts to consider the unique positions of the litigants before them and the specific nature of the proffered expert testimony before issuing an evidentiary ruling. The rules do not, on their face, create a blanket exclusions or admissions for categories of scientific evidence. Nevertheless, in practice, courts tend to repeatedly encounter the same research techniques, scientific paradigms, and categories of expert testimony.

---

[170] Michael Janofsky, *On Cigarettes, Health and Lawyers,* N.Y. TIMES (Dec. 6, 1993), https://www.nytimes.com/1993/12/06/business/on-cigarettes-health-and-lawyers.html.
[171] *Tobacco,* WORLD HEALTH ORGANIZATION (Jun. 25, 2025), https://www.who.int/news-room/fact-sheets/detail/tobacco; UNITED STATES OFFICE OF THE SURGEON GENERAL, THE HEALTH CONSEQUENCES OF SMOKING—50 YEARS OF PROGRESS: A REPORT OF THE SURGEON GENERAL (2014).

This can present the same methodological issues in different cases. Courts can and do then functionally adopt presumptive rules of inadmissibility, absent extenuating or unusual circumstances.

The clearest case of such a *de facto* blanket exclusion concerns the same basic issue that confronted the *Frye* court in 1923: polygraph ("lie detector") tests. The skepticism surrounding early lie detection technology foreshadows modern judicial treatment of certain categories of scientific evidence. Courts have often treated such techniques not merely with case-by-case scrutiny, but with a form of presumptive inadmissibility. In *United States v. Scheffer*, the Supreme Court upheld a flat rule that "makes polygraph evidence inadmissible in court-martial proceedings."[172] The Court reached this determination by finding that "there is simply no consensus that polygraph evidence is reliable."[173] To reach this conclusion, the Court reviewed studies which found that the accuracy rate of polygraph tests was somewhere in the range of 50–87%.[174] The Court also cited approvingly the fact that the Fourth Circuit Court of Appeals has adopted its own *per se* ban on polygraph evidence.[175]

These rationales also support a presumptive rule against the admissibility of animal model evidence. Multiple meta-analyses and systemic reviews have found that the accuracy rate of animal model evidence is much lower than 50–87%. As already described, 92% of all drugs that appear safe and effective in animal models fail in human clinical trials because they are either unsafe or ineffective.[176] Another retrospective review looked at all 101 therapies that

[172] United States v. Scheffer, 523 U.S. 303, 303 (1998).
[173] *Id.* at 309.
[174] *Id.* at 310.
[175] United States v. Sanchez, 118 F.3d 192, 197 (1997).
[176] U.S. FOOD AND DRUG ADMINISTRATION, INNOVATION/STAGNATION: CHALLENGE AND OPPORTUNITY ON THE CRITICAL PATH TO NEW MEDICAL PRODUCTS 8 (2004).

appeared "highly promising" in animal models that appeared in peer-reviewed publications between the years 1979 and 1983. Two decades later, only 27 of these therapies ever advanced to clinical trials. Of those 27, only 5 were approved for human use, and only *one*, remained in actual clinical use at the time the review was completed.[177]

And multiple trial courts that have been faced with making a ruling on the admissibility of animal model evidence have concluded that this type of evidence is too unreliable for Rule 702's demands.[178] One such ruling was handed down by Chief Judge Jack Weinstein in the Eastern District of New York. Chief Judge Weinstein presided over many significant mass "toxic tort" cases, and "produced a prodigious body of academic literature describing and analyzing the problems related to complex mass tort litigation."[179] When confronted with a case involving the herbicide known as "Agent Orange," he ruled that plaintiffs could not establish a causal connection between this chemical and adverse health outcomes in humans through admissible evidence, because their case rested largely on animal model evidence:

> The many studies on animal exposure to Agent Orange, even plaintiffs' expert concedes, are not persuasive in this lawsuit. In a jointly-authored article, Dr. Ellen K. Silbergeld writes that "laboratory animal studies are generally viewed with more suspicion than epidemiological studies, because they require making the assumption that chemicals behave similarly in different species."

---

[177] Despina G. Contopoulos-Ioannidis et al., *Translation of Highly Promising Basic Science Research Into Clinical Applications,* 114 AM. J. MED. 477, 477 (2003).

[178] *See, e.g.,* Wade-Greaux v. Whitehall Laboratories Inc., 874 F. Supp. 1441, 1453, 1480 (D.V.I. 1994) ("Sugar and table salt have been shown to [cause birth defects] in some animal species . . . The notion that one can accurately extrapolate from animal data to humans to prove causation without supportive positive epidemiologic studies is scientifically invalid because it is inconsistent with several universally accepted and tested scientific principles. The principle of species specificity has been tested and demonstrates that different species can react differently to the same agent"); Tyler ex. rel. Tyler v. Sterling Drug, Inc., 19 F. Supp. 2d 1239, 1244 (N.D. Okla. 1998) ("Test results on animals are not necessarily reliable evidence of the same reaction in humans"); In re Human Tissue Products Liability Litigation, 582 F. Supp. 2d 644, 670 (D.N.J. 2008) ("Courts have been reluctant to allow expert testimony based upon animal studies to prove causation in humans unless there are good grounds to extrapolate data and results from animals to humans").

[179] Linda S. Mullenix, *Individual Justice in Mass Tort Litigation: The Effect of Class Actions, Consolidations, and Other Multiparty Devices,* 5 L. & POL. BOOK REV. 173 (1995).

. . .

> The animal studies are not helpful in the instant case because they involve different biological species. They are of so little probative force and are so potentially misleading as to be inadmissible. See Fed. R. Ev. 401–403. They cannot be an acceptable predicate for an opinion under Rule 703.[180]

## VII. Conclusion

In 2023, Rule 702 was amended to emphasize that the proponent of the expert testimony has the burden to demonstrate to the court "that it is more likely than not that the proffered testimony meets the admissibility requirements set forth in the rule."[181] These requirements include that such expert testimony be "the product of reliable principles and methods," and that the witness demonstrate that they have reliably applied a reliable methodology.[182] In nearly all cases, the use of animal model evidence fails to satisfy this requirement, and the proponent of the expert testimony will be unable to meet this burden. Just as courts have come to treat polygraph evidence as presumptively unreliable, the demonstrated limitations of animal models justify a similar approach. Absent strong corroborating human evidence, animal model data cannot form the basis of admissible expert testimony under Rule 702.

Decades before *Daubert* transformed the test for admission of expert testimony in federal courts, the courts recognized that the "gatekeeper" requirement was important because witnesses who are presented to the jury as "experts" have an "aura of special reliability and trustworthiness."[183] Indeed, opinion polling appears to show that the general public favors the

---

[180] *In re* Agent Orange Product Liability Litigation, 611 F. Supp. 1223, 1241 (E.D.N.Y. 1985).
[181] FED. R. EVID. 702 advisory committee's note to 2023 amendment.
[182] FED. R. EVID. 702(c), (d).
[183] United States v. Amaral, 488 F.2d 1148, 1152 (9th Cir. 1973). *See also* United States v. Brown, 557 F.2d 541, 556 (6th Cir. 1977) (recognizing "the inherent danger that expert testimony admitted without proper foundation may tend to confuse or mislead the trier of fact and thus defeat a defendant's right to a fair trial"); Viterbo v. Dow Chem. Co., 646 F. Supp. 1420, 1426 (E.D. Tex. 1986) (excluding testimony from two proffered expert witnesses on the grounds

use of animals in research on their own expectation that Jurors will then give that testimony significant weight, even if it is based on unreliable methods or an unreliable application of those methods.[184] Such pseudoscience testimony lacks any foundation in sound empirical methods but nevertheless has the prestigious imprimatur of 'science,' leading to verdicts and judgments that aren't rooted in fact.

Because this veneer of expertise is detached from sound evidence-based principles, but can nevertheless have a powerful persuasive effect on a jury of laypersons, such evidence should also be excluded under Rule 403.[185] This Rule calls for the exclusion of evidence if its probative value is substantially outweighed by a danger of unfair prejudice, confusing the issues, or misleading the jury. There also exist persuasive policy reasons for strict adherence to Rule 702. Knowing that unreliable animal model evidence may nevertheless gain currency in federal courts, companies and other entities may continue to carry out unnecessary and pointless experiments on animals for the purpose of generating data to mislead juries and protect the company from potential liability.

For these reasons, the admission of animal model evidence warrants stricter judicial scrutiny. Under Rules 702 and 403, courts should adopt a stance of presumptive inadmissibility, for the same reasons that courts have done so with respect to polygraph evidence for more than a century. Such evidence is unreliable, leads to unjust and unsupported outcomes, and confuses the crucial truth-seeking functions of the rules of evidence.

---

that there was "a great possibility of misleading the jury through the creation of a false aura of scientific infallibility through the use of such testimony"), *aff'd*, 826 F.2d 420 (5th Cir. 1987).

[184] DEBORAH JONES MERRITT & RIC SIMMONS, LEARNING EVIDENCE 785–86 (5th ed. 2022).

[185] United States v. Rincon, 28 F.3d 921, 923 (9th Cir. 1994) ("The Court [in *Daubert*] stated that evidence otherwise admissible may be excluded under Rule 403 if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury.").