IF THE GLOVE DON’T FIT, TRY NEWER GLOVES: THE UNPLANNED OBSOLESCEENCE OF THE SUBSTANTIAL SIMILARITY STANDARD FOR EXPERIMENTAL EVIDENCE

Jonathan M Hoffman

Available at: https://works.bepress.com/jonathan_hoffman/1/
If the Glove Don’t Fit, Try Newer Gloves:
The Unplanned Obsolescence of the Substantial Similarity Standard
for Experimental Evidence

Jonathan M. Hoffman

ABSTRACT

In the context of a recent Fifth Circuit decision, this article reviews the law concerning the admissibility of “experimental” and demonstrative evidence. The standards used to determine the admissibility of both categories of evidence predate the Federal Rules of Evidence. These standards for admission of such evidence are obsolete and at odds with the Federal Rules. The issue is particularly important in the wake of the Kumho Tire decision and the 2000 amendments to Federal Rule of Evidence 702, as engineers and other technical experts are increasingly called upon to test their hypotheses, even as the courts’ continued use of the old standard for admissibility results in unpredictable rulings on the admissibility of evidence derived from such testing. This article advocates that the old common-law “substantial similarity” standard should be discarded in favor of rigorous application of the Federal Rules of Evidence.

INTRODUCTION

In the wake of Kumho Tire Co. Ltd. v. Carmichael, more experts than ever in scientific and technical fields are testing the theories about which they testify in major litigation. The admissibility of testing or “experimental” evidence, as well as the proper use of demonstrative evidence of such testing, is often crucial. The stakes are high. The cost of conducting proper testing can be considerable; the cost of losing the case if a party fails to test (or if her test is excluded) is enormous. The jury’s verdict may rest, in large measure, on demonstrative evidence

1 Jonathan M. Hoffman is a senior partner of the firm of Martin Bischoff Templeton Langslet & Hoffman, LLP, in Portland, Oregon. He received his A.B. from Harvard College magna cum laude and his J.D. from the University of Oregon School of Law. The author gratefully acknowledges the assistance of Prof. Laird Kirkpatrick for his insights and critiques. Any errors in the analysis contained in this article are the author’s own.
offered to persuade jurors by helping them understand the testing performed by the experts and its significance to the issues in dispute. If the court excludes the testing evidence in whole or in part, the proponent’s case may be doomed. The stakes are equally great for the opponent of such evidence. If the stakes are higher than ever before, so is the frequency with which the issue arises. This is partly because of the effect of *Kumho*, but it is also due to the increasing sophistication of forensic testing and demonstrative evidence.

Given its importance, one might expect that the rules of evidence would provide clear criteria governing the admissibility of experimental and demonstrative evidence. Unfortunately, this is not the case. The Fifth Circuit’s recent decision in *Muth v. Ford Motor Co.* illustrates persistent problems with federal courts’ treatment of experimental evidence. As a result of the *Kumho* decision and the subsequent revisions of FRE 702 experts now are typically required to test their hypotheses if possible, to do so in accordance with the rigors of their profession, and to apply the results reliably to the facts. But if evidence of tests which satisfies those requirements is offered at trial, courts continue to rule unpredictably on the admissibility of those tests. There are two principal reasons for this. First, courts continue to treat experimental evidence as a discrete category of evidence whose admissibility is judged by a nebulous substantial similarity standard which predates the Federal Rules and no longer serves a useful purpose. Second, the Federal Rules fail to provide clear guidance as to how demonstrative evidence should be handled.

---

3 461 F.3d 557 (5th Cir. 2006).
4 Rule 702 states:

**Testimony by Experts.**

If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.
Rigorous application of the Federal Rules would ensure that experimental evidence is reliable and relevant, as required by Rules 702 and 401, and that its relevance is not substantially outweighed by risks of unfair prejudice, as required by Rule 403. As long as courts respect the constraints that now exist under these, and other, Federal Rules, the imprecise substantial similarity standard from the past is obsolete. Continued use of the substantial similarity standard injects needless confusion and imprecision to evidentiary decisionmaking. Although similarities and dissimilarities between test conditions and accident conditions are a pertinent factor in assessing the reliability, relevance, and prejudicial impact of tests conducted by experts, the old substantial similarity standard for determining the admissibility of such evidence distracts litigants and courts from more specific criteria that ought to be explicitly considered. There is no longer any need to treat experimental evidence as a discrete category of evidence law governed by antiquated standards of admissibility.

As for demonstrative evidence, Rule 401 either ought to be understood to provide, or amended to make explicit, that “relevant evidence” includes evidence that fairly and accurately explains, illustrates, or clarifies other admissible evidence. With such a clarification, Rule 403 and the other Federal Rules would provide a suitable framework for judging the admissibility of demonstrative evidence that flows from testing than the old substantial similarity standard.

DISCUSSION

A. The Muth Decision

---

5 F.R.E. 403 provides:

Exclusion of Relevant Evidence on Grounds of Prejudice, Confusion, or Waste of Time.
Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.
Muth v. Ford Motor Co., demonstrates the shortcomings of the outdated substantial similarity standard. In Muth, an automobile passenger sustained a spinal cord injury resulting in quadriplegia in a rollover of the Ford Crown Victoria in which he was riding. Plaintiff brought a product liability claim against Ford, contending that the vehicle’s roof was insufficiently strong, and that plaintiff was injured as a result of the roof crush. Ford asserted that a stronger roof would not prevent head and neck injuries in rollover accidents because, during the rollover event, even a restrained occupant’s head drops as much as five inches toward the ground and that the spinal cord injury occurs before the roof crush occurs. Ford offered demonstrative evidence from both the Malibu tests and the Controlled Rollover Impact System (CRIS) test, to illustrate Ford’s contention that a stronger roof would do little, if anything, to prevent injuries in rollover accidents.

To date, no one has developed a dynamic rollover test that is both repeatable and capable of simulating real-world rollover deformation patterns. This is a major reason NHTSA’s regulatory standard continues to specify a static test for roof strength. Repeatability is difficult because a motor vehicle rolling over is somewhat analogous to tossing a football onto the ground. The Malibu and Controlled Rollover Impact System (CRIS) tests are two of the foremost attempts to develop dynamic tests for assessing the kinematics of rollover accidents.

General Motors designed the Malibu tests in the early 1980s. GM conducted a series of

---

6 Supra, n. 3.
7 Id., at 560.
8 Id.
9 Id., at 560-561.
10 Id., at 561.

Page 4 of 39
full-scale rollover crash tests at the General Motors Proving Grounds using instrumented anthropomorphic dummies, without restraints, in 1983 Chevrolet Malibus. Test films depicted a series of Malibu sedans on a “cradle” being pushed sideways down a track. The cars were “launched” when they reached the end of the track at a speed of 32 miles per hour. High-speed photos show the roll sequence from various vantage points, and also the two front seat dummies from a camera mounted in the back seat. The tests measured loads imposed on instrumented dummies, and compared the results between production vehicles and vehicles modified to make the roof rigid.

The CRIS test, conducted by Exponent, Inc., during 2000-2001, was an improved test developed in response to the National Highway and Traffic Administration’s (“NHTSA”) request for comments on Federal Motor Vehicle Safety Standard 216, which establishes a static strength standard for automobile roofs. The CRIS test improved on the Malibu tests by controlling the position, momentum, and point of impact of the vehicle’s first contact with the ground. NHTSA did not adopt it (or any other dynamic test protocol) when it last revised the federal safety standard in 2001, because the vehicle’s motion after first contact with the ground was not sufficiently repeatable; nevertheless, the agency concluded that the CRIS test was “helpful in understanding occupant kinematics during rollover crashes.”

(October 9-11, 1985).

13 Muth, at 561; Harvey v. General Motors Corp., 873 F.2d 1343, 1355 (10th Cir. 1989).
14 Harvey, supra, at 1355.
15 Id.
16 Id.
18 Muth, supra, 461 F. 3rd, at 566.
19 Supra, n. 11.
In *Muth*, Ford offered both tests to show that spinal cord injuries in rollovers are not caused by roof crush and to assist the jury in understanding Ford's experts’ testimony regarding the general dynamic characteristics of rollover accidents.\(^2\) The trial court permitted the experts to testify but excluded demonstrative evidence of the tests, including film and photographs, noting that the tests were not conducted under substantially the same conditions as those involved in the lawsuit.\(^2\)

The Fifth Circuit affirmed. The court recited the rule that when demonstrative evidence is offered only as an illustration of general scientific principles, not as a re-enactment of disputed events, it need not meet a substantial similarity standard.\(^2\) It cautioned that such demonstrative evidence must not be misleading in and of itself.\(^2\) A demonstration might mislead when it resembles the disputed accident, but, the court observed, this resemblance gives rise to the requirement of substantial similarity.\(^2\) The Fifth Circuit ruled that the trial court had the discretion to determine that the demonstrative evidence was not quite similar enough, yet too closely resembled, the disputed accident, to present abstract principles effectively without misleading the jury.\(^2\)

Ford claimed that the CRIS test was not offered as a re-enactment but only to show general scientific principles.\(^2\) The conditions of the CRIS test closely resembled Ford’s theory of the accident as presented by its reconstructionist but were sharply at odds with the

\(^2\) 461 F.3d 557 (5th Cir. 2006).
\(^2\) Id.
\(^2\) See Gilbert v. Cosco, Inc., 989 F.2d 399, 402 (10th Cir. 1993); Four Corners Helicopters, Inc. v. Turbomecca, S.A., 979 F.2d 1434, 1441-1442 (10th Cir. 1992); Swajian v. General Motors Corp., 916 F.2d 31, 36 (1st Cir. 1990) (disallowing admission of videotaped tests because defendant “fail[ed] to mention even one ... general scientific principle that the videotapes purport [ed] to demonstrate”).
\(^2\) Muth, 461 F. 3rd, at 566.
\(^2\) Id.
\(^2\) Id.
\(^2\) Id.
reconstruction evidence offered by plaintiffs. Given this fact, the trial court exercised its discretion to reject the video as not being merely a presentation of general scientific principles while permitting Ford’s expert witnesses to testify at length to their conclusions, limiting only their reliance upon “visual aids.” Accordingly, the Fifth Circuit affirmed the trial court’s exclusion of the test videos and photographs.

**B. Origins of the substantial similarity standard**

There is no indication that the parties or the court in *Muth* ever questioned the rationale for applying a substantial similarity standard to determine the admissibility of demonstrative evidence of tests that “re-enact” the accident. Such a standard is found nowhere in the Federal Rules or their comments. It predates *Daubert* and, indeed, the Federal Rules of Evidence. It should therefore come as no surprise that the rule’s origins are unrelated to existing standards for admissibility of expert testimony. Rather, well before the codification of evidence law, common-law decisions characterized “experimental evidence” as a discrete category and used a substantial similarity standard to judge the admissibility of such evidence.

1. **Pre-Federal Rules Cases.**

In one early (1917) example of “experimental evidence,” *May Dept. Stores Co. v. Runge*, plaintiff sustained an on-the-job injury when a 200-pound “truck” fell down an elevator shaft and landed on him. The principal issue at trial was whether the truck had somehow passed through a gate on an upper floor of the building. The parties disputed whether there had been sufficient clearance below the gate for the truck to pass through. The gate was raised and

---

27 *Id.* at 567.
28 *Id.*
29 *Id.*, at 566.
lowered by a rope. Plaintiff claimed that the rope had been broken and retied several weeks
before the accident and had thereby been made so short that it would not allow the gate to
descend sufficiently to block access to the elevator shaft, and that the rope and the gate had been
in this condition for at least two or three weeks prior to the accident.\(^{32}\)

Defendant disputed this, claiming that the truck was probably in the elevator shaft when
plaintiff used the elevator. It offered the testimony of fact witnesses who, after the accident,
attempted to raise the gate as high as it could go.\(^{33}\) They tied the rope so that the gate would be
raised as high as possible, consistent with the automatic operation of the elevator. Then they
tried to push a truck of the same dimensions and just like that which injured plaintiff under the
gate into the elevator shaft. They found that it was impossible to do so.\(^{34}\) The trial court rejected
defendant’s evidence as “self-serving,” but the Eighth Circuit reversed. It concluded that the
evidence was no more self-serving than other evidence offered by plaintiff, and that it was
admissible as experimental evidence:

If the question were whether or not at a time past an article of furniture passed
through a specific door or window, whether or not a ladder reached the top of a
pile of lumber or of a house, whether or not a rope was long enough to reach
between two fixed objects, and the testimony of witnesses was conflicting,
subsequent actual tests would be competent, relevant, and convincing evidence of
the fact. To the common mind such evidence is even more persuasive and
convincing than the testimony of witnesses who casually looked at the articles in
some past time.\(^{35}\)

Citing cases going back to 1872,\(^{36}\) the court concluded that such evidence was admissible if it
“substantially tends to establish the fact it is offered to prove.” \textit{May Stores} cited older cases

\(^{31}\) 241 F. 575 (8th Cir. 1917).

\(^{32}\) \textit{Id.}, at 577.

\(^{33}\) \textit{Id.}

\(^{34}\) \textit{Id.}, at 577, 578.

\(^{35}\) \textit{Id.} at 578.

holding that tests could be competent evidence if the conditions of the tests were the same as those of the plaintiff’s accident “as near as practicable.”37 Where it is “doubtful whether or not it has such a tendency on account of its remoteness in time, place, or otherwise, and where it is likely to tend more to confusion and inconvenience than to justice and certainty, it is discretionary with the trial court to limit the extent to which it may be received.”38 Notably, like many of the early cases, May Stores had nothing to do with expert witness testimony, as the witnesses who conducted the tests were not experts.

Other cases continued to apply a similarly substantial standard for the admission of experimental evidence, without focusing on whether the evidence was presented through fact witnesses or experts.39 More recently, but still decades before adoption of the Federal Rules, the Third Circuit applied such a standard to decide the admissibility of an expert’s test to determine the cause of an explosion in an oil well, holding that the test conditions were sufficiently similar to allow a logically relevant inference.40

In each of the above cases, courts applied the standard to evidence of the tests themselves, as contrasted with the Fifth Circuit’s application of the standard in Muth solely to demonstrative evidence of the tests. However, other pre-F.R.E. cases applied the substantial similarity standard to in-court experiments or demonstrations, and employed the same rationale as for out-of-court tests.41


38 Id. at 578.
40 Hopkins v. E. I. du Pont De Nemours & Co., 199 F.2d 930, 934 (3rd Cir. 1952).
Following adoption of the Federal Rules, some courts continued to treat “experimental evidence” as a discrete category of evidence, without questioning whether the earlier cases were consistent with the Federal Rules. The Tenth Circuit’s decision in *Harvey v. General Motors*, for example, addressed the admissibility of the same Malibu tests later at issue in *Muth*, by continuing to apply the pre-FRE “experimental evidence” standards:

Where experiments such as this are not based on the facts, however, it must be made clear to the jury that the evidence is admitted for a limited purpose.... Where, however, an experiment purports to simulate actual events and to show the jury what presumably occurred at the scene of the accident, the party introducing the evidence has a burden of demonstrating substantial similarity of conditions. They may not be identical but they ought to be sufficiently similar so as to provide a fair comparison.

In *Fusco v. General Motors Corp.*, plaintiff claimed that the vehicle lost control and struck a tree because of fatigue failure of a ball stud in the front suspension. GM countered that the ball stud failure was a result, not a cause, of the collision with the tree. Its expert conducted a videotaped test to show that if the stud had broken before the car veered, there would have been a heavy black tire mark on the road because the uncontrolled tire would have dragged as the car slid off course. The trial judge excluded the test and the First Circuit affirmed. It characterized the issue as “interesting and important,” and acknowledged that the case law surrounding such evidence was “muddled.” It observed:

The concern lies not with use of tape or film (the issue would be largely the same if the jurors were taken to the test track for a live demonstration) but with the deliberate recreation of an event under staged conditions. Where that recreation could easily seem to resemble the actual occurrence, courts have feared that the jurors may be misled because they do not fully appreciate how variations in the surrounding conditions, as between the original occurrence and the staged event,

---

42 See, e.g., *McKnight v. Johnson Controls, Inc.*, 36 F.3rd 1396, 1402 (8th Cir. 1994).
43 873 F.2d 1343 (10th Cir. 1989).
44 *Id.*, at 1356. Curiously, the court in *Muth* did not cite *Harvey*, even though it addressed the admissibility of one of the two tests excluded in *Muth*.
45 11 F.3d 259 (1st Cir. 1993). *cited in Muth, supra*, 461 F. 3rd, n. 20 at 566.
can alter the outcome. In such cases, the solution of many courts, including this one, has been to call for substantial similarity in conditions, or to stress the great discretion of the trial judge to exclude the evidence where similarity is not shown, or both.\(^{46}\)

The First Circuit characterized this rule as a doctrine, “predating and now loosely appended to Rule 403.”\(^{47}\) However, the court did not explore the relationship between Rule 403 and the substantial similarity standard, tacitly treating them as analogous, if not identical. Had it looked further, it might have recognized significant differences, even inconsistencies, between the two.\(^{48}\) The First Circuit’s opinion also attempted to link the substantial similarity standard to the Supreme Court’s decision in *Daubert*, because Rule 702 imposed a reliability requirement for expert testimony.\(^{49}\) Although not articulated by the court, the implication of its comments is that a re-creation of the accident that was not substantially similar could be excluded as unreliable.

Historically, many in-court “experiments” were no doubt less elaborate and scientific than those at issue in *Fusco, Harvey* and *Muth*, and did not necessarily involve experts, or even witnesses. Some were simply performed by the lawyers and may have been more akin to Perry Mason-style in-courtroom “stunts.” Wright & Graham’s treatise chronicles some real-life extreme examples of these kinds of courtroom “experiments” which serve “no useful analytic purpose, [yet] may still function as an excuse for trial judges to take leave of their senses and attorneys to stretch the ethics of advocacy to the breaking point.”\(^{50}\) Perhaps the last such notorious in-court demonstrative experiment was the most disastrous: the Assistant District Attorney’s in-court experiment with the glove that didn’t fit during the O.J. Simpson trial.

Many of the reported cases arose from attempts by parties to “re-create” the accident or

\(^{46}\) *Id.*, at 263-264.

\(^{47}\) 11 F. 3rd, at 264. The text of Rule 403 is shown in n. 5, *infra*.

\(^{48}\) See discussion, *infra*, at pp. -----.

\(^{49}\) *Id.*
incident at issue in the case, with varying degrees of verisimilitude. Historically, accident
reconstructions or re-creations were viewed with suspicion, and expert testimony to reconstruct
or re-create an accident was itself viewed with substantial skepticism, if not outright mistrust.
This was because such evidence was variously considered conjectural, speculative, misleading,
and because it could invade the province of the jury in resolving disputed facts. \footnote{Logsdon v. Baker, 366 F.Supp. 332 (D.D.C. 1973), vacated and remanded, 517 F.2d 174 (D.C. Cir. 1975); see also, Admissibility of Opinion Evidence as to Point of Impact or Collision in Motor Vehicle Accident Case, 66 A.L.R.2d 1048 (1959); Burns, The Role of Reconstruction Experts in Witnessed Accident Litigation, 22 De Paul L.J. 7 (1972).}
Experimental evidence was also used for a variety of other purposes, such as to demonstrate that something posited by the proponent was possible or that something asserted by the adversary was impossible; to challenge the non-physical abilities of witnesses (e.g., to be able to see, hear, or smell) \footnote{Newman, who turns into a criminal defense lawyer to represent a friend in a major murder case and wins an acquittal by causing a key witness to fail a test of his ability to differentiate the smell of various liquors in the classic film, “The Young Philadelphians” (Warner Bros. 1959).}
; or to challenge a purportedly common-sense proposition. \footnote{Note, The Use of Ad Hoc Experimental Evidence in Litigation, 104 U. Pa. L. Rev. 1064 (1957).}
Even after the Federal Rules were adopted, some courts continued to label such evidence as “experimental evidence” and drew no distinction between the substantive evidence and the manner in which it was presented. \footnote{See Champeau v. Fruehauf Corp, 814 F.2d 1271, 1278 (8th Cir. 1987) (discussing rules for admission of ‘experimental evidence’); see also, McKnight v. Johnson Controls, Inc., supra, Jodoin v. Toyota Motor Corp, 284 F.3d 272 (1st Cir. 2002); Chase v. General Motors Corp., 856 F.2d 17 (4th Cir. 1988); Gladhill v. General Motors Corp., 743 F.2d 1049 (4th Cir.1984). See also, cases listed and discussed in Robert E. Jones & Gerald E. Rosen,}

\section*{C. Demonstrative Evidence–Whatever It Is.}

\footnotetext[50]{22 Wright & Graham, Fed. Prac. & Proc. § 5172, at 124-126 (1978).}
\footnotetext[52]{An illustrative example from popular culture is the ambitious young tax lawyer portrayed by a youthful Paul Newman, who turns into a criminal defense lawyer to represent a friend in a major murder case and wins an acquittal by causing a key witness to fail a test of his ability to differentiate the smell of various liquors in the classic film, “The Young Philadelphians” (Warner Bros. 1959).}
\footnotetext[53]{Note, The Use of Ad Hoc Experimental Evidence in Litigation, 104 U. Pa. L. Rev. 1064 (1957).}
\footnotetext[54]{See Champeau v. Fruehauf Corp, 814 F.2d 1271, 1278 (8th Cir. 1987) (discussing rules for admission of ‘experimental evidence’); see also, McKnight v. Johnson Controls, Inc., supra, Jodoin v. Toyota Motor Corp, 284 F.3d 272 (1st Cir. 2002); Chase v. General Motors Corp., 856 F.2d 17 (4th Cir. 1988); Gladhill v. General Motors Corp., 743 F.2d 1049 (4th Cir.1984). See also, cases listed and discussed in Robert E. Jones & Gerald E. Rosen,
In *Muth*, the trial court excluded demonstrative evidence of the Malibu and CRIS tests but permitted other evidence of the tests. In effect, the court treated the admissibility threshold for demonstrative evidence as more stringent than that for other evidence. Conversely, however, one commentator suggests that demonstrative evidence is permitted to be used more liberally than other evidence because it is shown to the jury but not formally admitted as evidence and therefore is not required to meet the higher post-*Daubert* admissibility standards.  

Inconsistent treatment of demonstrative evidence should come as no surprise. Despite the importance of such evidence in complex modern trials and the frequency with which it is used, the Federal Rules of Evidence contain neither a definition of demonstrative evidence nor any rule spelling out how a trial court should determine its admissibility or other use at trial. The only mention of the subject is a passing reference contained in the Advisory Committee Note to Rule 611(a). Rule 611(a) gives the court wide discretion over the mode and order of presentation of evidence at trial. However, neither the rule nor the Advisory Committee Note gives any concrete guidance over how that discretion should be exercised, particularly with respect to demonstrative evidence. Indeed, the Advisory Committee Note creates the potential for

---

_Federal Civil Trials and Evidence, §§ 8:660-8:706, at pp. 8c-75-81 (2006)._  
57. *Id.* n. 10 at 961, 1017. The Advisory Committee Note to Rule 611(a) states that the first subdivision of the rule “covers such concerns as whether testimony shall be in the form of a free narrative or responses to specific questions, McCormick § 5, the order of calling witnesses and presenting evidence, 6 Wigmore § 1867, the use of demonstrative evidence, McCormick § 179, and the many other questions arising during the course of a trial which can be solved only by the judge's common sense and fairness in view of the particular circumstances.”  
58. Rule 611(a) states:  

**Control by court.**  
The court shall exercise reasonable control over the mode and order of interrogating witnesses and presenting evidence so as to (1) make the interrogation and presentation effective for the ascertainment of the truth, (2) avoid needless consumption of time, and (3) protect witnesses from
additional confusion because of the absence of a cross-reference to any other rule that might bear upon the use of demonstrative evidence.59

Most of the leading publications on the subject of demonstrative evidence were authored by prominent practitioners, most notably Melvin Belli, rather than legal scholars,60 and a “concomitant avalanche” of others as well.61 Most such writings popularized demonstrative evidence rather than analyzed it, and failed to present any unifying theory explaining its relevance, admissibility, or use.62 One book seriously discussed the evidentiary rules with respect to various types of demonstrative evidence.63 But it said little more than that the admission of demonstrative evidence is discretionary, and that its use is limited by Rule 403. This is not a criticism of the book but rather is indicative of the lack of clear guidance either in the Federal Rules of Evidence or the case law.

Scholarly writings are of little additional help. Brain & Broderick’s excellent article highlights many of the problems with the publications of their forebears. They note the absence of any scholarly theory explaining the relevance of demonstrative evidence,64 as well as some of the definitional and conceptual shortcomings of the leading scholarly works on the subject.65 Unfortunately, their article predates Daubert and the subsequent revolution in expert evidence. However, they do acknowledge the “ever-increasing emphasis in modern trials on lay and expert

59 The Advisory Committee Notes to Rule 611(a) do cross-reference Rule 403(b), but only to the extent of the discretion vested in the judge to exclude evidence “as a waste of time in Rule 403(b).”
60 See, e.g., MELVIN BELLI, MODERN TRIALS (1954) and READY FOR THE PLAINTIFF! (1956).
61 See sources cited in BRAIN & BRODERICK, at 999, n. 145 at 1000.
62 Id., at 1002.
64 BRAIN & BRODERICK, at 959.
65 Id., at 1004-1012.
opinion testimony” and the resulting fact that “modern lawyers are increasingly turning to demonstrative evidence to make these opinions understandable to triers of fact.” They propose a cogent definition of demonstrative evidence as that which “is principally used to illustrate or explain other testimonial, documentary, or real proof, or a judicially noticed fact. It is, in short, a visual (or other sensory) aid.” They point out that a major shortcoming of the existing rules of evidence is a failure to provide an explicit avenue for the admissibility of demonstrative evidence. Indeed, strictly speaking, such demonstrative evidence ordinarily does not prove or disprove facts of consequence and therefore cannot meet the definition of “relevant evidence” under Rule 401. Therefore, they advocate modifying Federal Rule 401 to provide additional guidance to the admissibility of demonstrative evidence by defining “relevant evidence” to include: (a) evidence having any tendency to make the apparent existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence. . .; or (b) evidence that fairly and accurately explains, illustrates, or clarifies other admissible evidence.”


To some extent, the substantial similarity standard may have served a function roughly analogous to, albeit less systematic than, the reliability and relevance standards of the Federal Rules today. After all, some older cases justified the substantial similarity standard in a manner closely tracking the relevancy provisions of Rules 401 and 403. Although the Federal Rules do not explicitly define or address either experimental or demonstrative evidence, some courts have

---

66 Id., at 964.
67 Id., at 968-969.
68 Id., at 1023.
69 See e.g., n. 38-39, supra.
applied constraints contained in the existing Federal Rules to curtail the inappropriate use of tests or experiments as demonstrative evidence by experts.\textsuperscript{70}

\section*{1. Rule 702.} \textsuperscript{71}

Rule 702 requires that any such evidence, whether testimonial or demonstrative, is reliably obtained and applied by qualified experts, is based upon sufficient facts or data, that the testimony is the product of reliable principles and methods, and that the witness has applied the principles and methods reliably to the facts of the case. In determining reliability under Federal Rule 702, the Court in \textit{Daubert}\textsuperscript{72} abandoned the \textit{Frye} test,\textsuperscript{73} a pre-Federal Rules standard for the admissibility of novel scientific evidence, replacing \textit{Frye} with a nonexclusive list of key factors for courts to consider in determining the reliability of expert testimony including whether the theory or technique can be and has been tested.\textsuperscript{74}

\textit{Daubert} was explicit that “\textit{[o]ur discussion is limited to the scientific context because that is the nature of the expertise offered here.}”\textsuperscript{75} In the following decade, lower courts reached differing conclusions as to whether those criteria applied solely to scientific testimony, solely to \textit{novel} scientific testimony, or to \textit{all} expert testimony.\textsuperscript{76} However, a few courts quickly concluded that many of the same hallmarks of reliability should also extend to the testimony of technical

\begin{thebibliography}{9}
\bibitem{70} It should be noted that the evidence codes of many states have diverged from the Federal Rules by failing to adopt the 2000 Amendments to Rules 701, 702, and 703. As a result, state evidence law may not provide some of the important and meaningful constraints that, to this writer, now render the substantial similarity standard superfluous under the Federal Rules. Accordingly, this article is not intended to suggest, one way or the other, whether the substantial similarity standard may still serve a useful purpose under state evidence law.
\bibitem{71} See n. 4, supra.
\bibitem{72} 509 U.S. 579 (1993).
\bibitem{73} \textit{Frye v. United States}, 293 Fed. 1013 (D.C. Cir. 1923).
\bibitem{74} 509 U.S. at 592-594.
\bibitem{75} 509 U.S. n. 8, at 590. See also Chief Justice Rehnquist’s concurrence: “Does all this \textit{dicta} apply to an expert seeking to testify on the basis of ‘technical or other specialized knowledge’--the other types of expert knowledge to which Rule 702 applies--or are the ‘general observations’ limited only to ‘scientific knowledge?’” 509 U.S. at 601.
\bibitem{76} Compare, e.g., Compton v. Subaru of America, 82 F.3rd 1513 (10th Cir.), \textit{cert. denied}, 519 U.S. 1042
\end{thebibliography}
experts. A forceful early example was *Stanczyk v. Black & Decker, Inc.*,\(^{77}\) in which the court excluded the testimony of an engineer who failed to test his theory about the utility of his proposed guard for a miter saw. Judge Zagel illustrated the issue by contrasting the theories of Seventeenth-Century astronomers Tycho Brahe and Galileo—the former theorizing without attempting to confirm his theories through observation, and the latter testing and refuting long-held beliefs through empirical testing. Judge Zagel thereupon posed the evidentiary question thus: “The question before me is whether plaintiff's expert is a modern version of Tycho Brahe and, if so, does that mean his testimony ought to be excluded under *Daubert* . . .”?\(^{78}\)

A series of subsequent Seventh Circuit cases held that, given the importance of testing in the design process of certain kinds of products, an engineer’s failure to test a proposed alternative design rendered his opinion inadmissible.\(^{79}\) In another case, the court concluded that the admissibility of an engineer’s expert opinions should be based on considerations such as, “what tests do engineers use to resolve questions of the kind [this expert] addressed? What tests should he have performed? What data did he overlook?”\(^{80}\) Professor Faigman correctly predicted the ultimate outcome of the issue, stating that “[T]he guiding principle . . . must be *Daubert*’s expectation that hypotheses that can be tested will be tested.”\(^{81}\)

The Supreme Court finally resolved this question in *Kumho*, holding that *Daubert* requires a "fit" between the expert's reasoning and conclusions, and applies to all technical or

---


\(^{79}\) *Cummins v. Lyle Indust.*, 93 F.3d 362 (7th Cir. 1996).

\(^{80}\) *DePaepe v. General Motors Corp.*, 141 F.3rd 715, 719-720 (7th Cir.), cert. denied, 525 U.S. 1054 (1998).

other specialized expert testimony, not just scientific evidence. In 2000, Kumho’s holding was codified in an amended version of FRE 702, which provided that a qualified expert could testify to an opinion or otherwise, “if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.”

The effect of applying these reliability criteria to technical expert testimony has been to increase dramatically not only the quantity but also the quality of empirical testing by experts. In the years following the Kumho decision, lower courts generally began exercising greater scrutiny over the expert testimony of engineers and other technical experts. Although a few remained reluctant to require such testing in all cases in which the expert’s opinion was readily testable, the unmistakable trend has been to the contrary. A flurry of recent cases confirm that, as Professor Feigman asserted over a decade ago, when an expert’s opinion can be tested, it must be tested to be admissible.

However, just as Rule 702 requires experts to test their hypotheses if possible, it also

---

82 526 U.S. 147-149.
83 The Advisory Committee Notes to the 2000 amendments stated:

> The expert's testimony must be grounded in an accepted body of learning or experience in the expert's field, and the expert must explain how the conclusion is so grounded. See, e.g., American College of Trial Lawyers, Standards and Procedures for Determining the Admissibility of Expert Testimony after Daubert, 157 F.R.D. 571, 579 (1994) (“W)ether the testimony concerns economic principles, accounting standards, property valuation or other non-scientific subjects, it should be evaluated by reference to the 'knowledge and experience' of that particular field.”).
85 Id.
requires tests to be properly conducted and reliably applied to the facts of the case. Outside of litigation, engineers are required to make judgments about whether a test is adequately representative of a given set of facts. Likewise, in litigation, courts now apply Rule 702 to consider whether any dissimilarities between the test and the accident facts are sufficiently material to the issues in dispute to require exclusion of the expert’s testimony as unreliable. For example, in *Kumho’s* wake, a number of courts have critically assessed the reliability of the analysis of accident reconstructionists, and have excluded speculative or otherwise unreliable expert accident reconstruction evidence under Rule 702.88

2. **Rule 403.**89

Rule 403 gives trial courts broad discretion to exclude or restrict the use of relevant evidence that may be cumulative, misleading, or unfairly prejudicial. It does not prevent the use of evidence which is merely prejudicial; rather, it “protects against evidence that is unfairly prejudicial, that is, if it tends to suggest decision on an improper basis... commonly, though not necessarily, an emotional one.”90 From its inception, Rule 403 was suggested as a restriction on demonstrative evidence. Professor Schwartz, commenting on the proposed Federal Rules, remarked that the rule, then codified as Rule 4-03, might be interpreted as precluding “the more stimulating variety of real proof” and might therefore be a “cause for concern to advocates of the

---


89 See n. 5, supra.

benefits of demonstrative evidence." Likewise, DOMBROFF cites an early example in which a court applied Rule 403 to uphold the exclusion of demonstrative evidence: a chart showing the descent profile of an airliner that crashed on approach because the chart was misleading for failing to conform to the stipulated facts.92

Likewise, Rule 403 has been applied to exclude experimental evidence when tests performed by experts were found to be misleading.93 Indeed, two decades before it decided Muth, the same court addressed a similar issue involving the use of a crash test and decided the issue based on Rule 403.94 The court upheld the trial court’s exclusion of the film, after the lower court balanced the film's acknowledged relevance against the danger its admission would mislead, confuse, or prejudice the jury. The court determined that a limiting instruction would not defuse such improper influence, concluding that under Rule 403, the jury would likely consider it as more than a simple demonstration of general principles.95 The trial court had excluded rollover crash test films offered by both parties and took the consistent position that any such films using vehicles other than the subject model vehicle could be misleading.96

Surprisingly, the court in Muth did not cite the earlier Fifth Circuit case.

A number of other courts have applied Rule 403 to determine the admissibility of tests

92 In re Air Crash Disaster at John F. Kennedy Intern. Airport on June 24, 1975, 635 F.2d 67, 72 -73 (2nd Cir. 1980).
94 Shipp v. General Motors Corp., 750 F.2d 418, 427 (5th Cir. 1985).
95 Id. See also U.S. v. Gaskell, 985 F.2d 1056, 1062 (11th Cir. 1993) (“The ability to cross-examine is not a substitute for the offering party's burden of showing that a proffered demonstration or experiment offers a fair comparison to the contested events. Particularly where the demonstration unfairly tended to prejudice the jury on the one genuinely contested issue, without providing any significantly probative testimony, neither the cautionary instruction nor the ability to cross-examine was sufficient to cure the error.”).
96 Shipp, supra, 750 F.2d at 427 (“A central tenet of GM's argument was that all rollover accidents were different. Yet, its filmed ‘accident’ was a multiple, rather than a single revolution one (plaintiff’s rollover involved a single roll), involving a different vehicle with a substantially different roof and passenger compartment than the
offered to prove the party’s theory of the accident, where the proponent omitted important variables. For example, in *Chase v. General Motors Corp.*, plaintiff alleged that the crash occurred because of a defect in the braking system; GM countered that the crash was unavoidable. Plaintiff offered tests by NHTSA and by its own expert to show that, absent a defect, the brakes were capable of stopping the car in time to avoid the crash. However, the tests were performed in daylight, on flat, straight surfaces under controlled protocols at test facilities or on an airport runway. The tests used experienced drivers at regulated vehicle speeds and who steered straight ahead. Brake pedal force in the NHTSA tests was mechanically applied to produce single axle lock. The NHTSA tests were performed on three different surface types under wet conditions. In the expert’s tests, he applied the brakes himself, similarly tested the vehicle on both wet and dry surfaces.

However, the conditions on the night of the crash were significantly different from those in the tests. Plaintiff was driving downhill on a slight curve to the left and was so turning his car at the time of the accident. The surface of the road was covered with snow and ice. Instead of applying measured or pre-ascertained brake pressure, plaintiff instinctively applied the brakes. The Fourth Circuit analyzed the admissibility issues in a manner that conflated the

---

97 See, e.g., *Chase v. General Motors Corp.*, 856 F.2d 17 (4th Cir. 1988); *Gladhill v. General Motors Corp.*, 743 F.2d 1049 (4th Cir.1984).
98 856 F.2d 17 (4th Cir. 1988).
99 Id., at 18.
100 Id., at 18-19.
101 Id., at 19.
102 Id.
103 Id.
104 Id.
105 Id.
106 Id.
107 Id.
108 Id.
109 Id.
substantial similarity test with the language of Rule 403, stating that the conditions of the test “must be sufficiently close to those involved in the accident at issue to make the probative value of the demonstration outweigh its prejudicial effect.” 110 It further held that the trial judge should have excluded both evidence of the tests themselves, as well as the videotape of the tests. 111

In another case, plaintiff was injured when his vehicle struck a moose on the highway. 112 The district court excluded a video of an inverted drop test of a vehicle similar to the one in which plaintiff was injured. Plaintiff offered the test to show that the roof of the vehicle lacked sufficient strength. The court found that the test would be misleading to the jury and was therefore inadmissible under Rules 701, 702, and 403. 113 There was no evidence that the magnitude of the forces in the drop test and the moose collision were comparable to each other. 114 The forces imposed during the drop test were vertical, whereas the vehicle was moving forward when it struck the moose. 115 The test protocol indicated that it was designed to assess roof strength in rollovers, not animal collisions, and the test contained no performance standard by which to gauge its effectiveness. 116 After analyzing the published protocol for the test, the court excluded the tests because of “a high likelihood that the jury would be substantially affected by viewing the dramatic impact of the crush occurring as a vehicle is dropped on its roof from a height of 18 inches without gaining any significant understanding concerning what is a safe roof integrity against which to measure what happened” in plaintiff’s collision with the

110 Id., at 19, citing Gladhill, supra, 743 F.2d at 1051.
111 Id., at 20, citing Gladhill, supra, at 1052.
113 Id. at 190.
114 Id. at 188.
115 Id.
116 Id., at 189-190.
moose. Thus, the court was able to resolve the admissibility issue by reference to Rule 702 and 403, without resort to the substantial similarity standard. Instead, it properly focused on the specific factors that rendered the test unreliable, irrelevant to the issues in the case, misleading and unfairly inflammatory.

3. Rule 703.

Rule 703 permits an expert to render an opinion even if it is based on inadmissible evidence, as long as the evidence is “of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject.” This rule can be applied to impose constraints upon an expert’s demonstrative evidence in two circumstances. First, as a result of the 2000 amendments to Rule 703, it is now clear that if the expert reasonably relies on an inadmissible test to reach his or her opinion (such as one not personally conducted or witnessed by a testifying expert or any other witness), the test itself is presumptively inadmissible. Rule 703 permits the expert to render the opinion but does not permit the expert to bring the otherwise inadmissible data into evidence through the back door unless “the court determines that their probative value in assisting the jury to evaluate the expert's opinion substantially outweighs their prejudicial effect.” Second, if the court determines that the test was invalid, unreliable, or

---

117 Id. at 190.
118 Rule 703 provides:

**Bases of Opinion Testimony by Experts**

The facts or data in the particular case upon which an expert bases an opinion or inference may be those perceived by or made known to the expert at or before the hearing. If of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject, the facts or data need not be admissible in evidence in order for the opinion or inference to be admitted. Facts or data that are otherwise inadmissible shall not be disclosed to the jury by the proponent of the opinion or inference unless the court determines that their probative value in assisting the jury to evaluate the expert's opinion substantially outweighs their prejudicial effect.

119 See U.S. v. Leason, 453 F.3rd 631, 638 (4th Cir. 2006), cert. denied, 127 S.Ct. 1874 (2007), (upholding determination that probative value of hearsay relied upon by expert substantially outweighed prejudicial effect); Turner v. Burlington Northern Santa Fe R. Co., 338 F.3rd 1058, 1062 (9th Cir. 2003) (upholding exclusion of
otherwise was not “of a type reasonably relied upon by experts in the particular field,” it can preclude the expert’s opinion altogether, not simply the test.

In *Muth*, the court might have excluded evidence of the Malibu and CRIS tests under Rule 703 if the testifying expert did not participate in or have personal knowledge of the tests, but there is no indication that the parties or the court were concerned about demonstrative evidence of the tests being used as “backdoor hearsay” contrary to the limitations of Rule 703. The *Muth* opinion gives no indication that the parties contended, or that the court determined, that the tests were not “of a type reasonably relied upon” by experts. Indeed, the *Muth* opinion indicates that the experts were permitted to testify about the tests, and were merely precluded from showing any test films or photographs to the jury.

4. Rule 701.

Because so many of the recent cases arise from demonstrative evidence of tests performed by experts, it is easy to overlook the fact that experimental evidence can also be offered by lay witnesses as well. Indeed, many of the older cases arise from testimony of lay witnesses who “tested” a party’s theory of the accident. Therefore, even if Rule 702’s reliability requirements obviate the necessity of retaining the substantial similarity standard for experts, one might argue

---

120 Compare, *Montgomery v. Mitsubishi Motors Corp.*, 448 F. Supp. 2nd 619, 633 (E.D. Pa. 2006), (testifying expert could offer an opinion based on tests conducted by a non-testifying expert, but that this did not necessarily mean that evidence of the tests themselves would be admissible as evidence at trial); see also, *Mike’s Train House, Inc.*, v. *Lionel, L.L.C.*, 472 F.3d 398, 409 (6th Cir. 2006).

121 Rule 701 provides:

**Opinion Testimony by Lay Witnesses.**

If the witness is not testifying as an expert, the witness' testimony in the form of opinions or inferences is limited to those opinions or inferences which are (a) rationally based on the perception of the witness, and (b) helpful to a clear understanding of the witness' testimony or the determination of a fact in issue, and (c) not based on scientific, technical, or other specialized knowledge within the scope of Rule 702.

122 See, e.g., *May Dept. Stores*, discussed *supra*, at pp. 7-9.
that it is still necessary because of the risk that experiments by lay witnesses could be used to circumvent those requirements.

Whatever merit such a concern might have had in the past, it is superfluous today. The 2000 amendment to Rule 701 added the proviso that testimony offered under Rule 701 is “not based on scientific, technical, or other specialized knowledge within the scope of Rule 702.” This amendment was intended to make clear that the rule should not be used as a means to circumvent the requirements of Rule 702. Indeed, even before the 2000 amendments, one circuit had reached a similar conclusion under the earlier version of Rule 701.\textsuperscript{123} The commentary and case law surrounding Federal Rule 701 now protect against using lay witnesses to obtain admission of otherwise unreliable or prejudicial experiments or tests.\textsuperscript{124}

Furthermore, the “helpfulness” requirement of Rule 701(b) limits the improper use of tests and demonstrations by lay witnesses. If a jury can look at the same evidence considered by the lay witness and reach its own conclusions, the court may exclude as unhelpful the lay opinion of a witness not qualified as an expert.\textsuperscript{125}

E. Substantial Similarity: The Goldilocks Principle\textsuperscript{2}

Continued use of the substantial similarity standard for experimental evidence is no longer necessary in light of the foregoing Federal Rules, as they have been interpreted and applied. As already pointed out, Rule 702 requires experts to test their testable hypotheses, requires the court to act as a gatekeeper to ensure the reliability of the methods applied by the

\textsuperscript{123} Asplundh Mfg. Div. v. Benton Harbor Enr’g, 57 F.3d 1190, 1193 (3rd Cir. 1990).
expert, and requires the court to determine that the opinions are derived reliably from the facts. Rule 403 permits the court to exclude or limit otherwise relevant evidence if its relevance is substantially outweighed by the danger of unfair prejudice,\textsuperscript{126} confusion of the issues,\textsuperscript{127} or is misleading\textsuperscript{128} or cumulative,\textsuperscript{129} and Rule 703 permits the court to exclude backdoor hearsay upon which the expert relied. If any additional legitimate purpose is served by superimposing a separate substantial similarity test of reliable demonstrative evidence, the cases have failed to identify it. Instead, some courts, such as the court in \textit{Muth}, have continued to apply the substantial similarity standard, as though “experimental evidence” were still a discrete category of evidence in need of a special unwritten rule.

Even though the substantial similarity standard is an artifact of the pre-FRE case law, some might argue that it is a harmless one. After all, of what relevance is a courtroom test that is not substantially similar to the case at issue? As it turns out, there are plenty of reasons that continued use of such a standard is not a harmless error.

\begin{enumerate}
\item \textbf{1. Distraction From More Appropriate Standards.}
\end{enumerate}

The substantial similarity standard distracts the court from the analysis it ought to conduct under Rules 401, 403, 702, 703 and 105. It has led to less nuanced and more arbitrary exercises of the court’s discretion, such as a blanket rule to exclude all tests involving different model vehicles,\textsuperscript{130} or a decision to exclude all demonstrative testing evidence.\textsuperscript{131} Once a judge concludes that a test constitutes some kind of re-creation of the accident and makes a finding that the test is, or is not, substantially similar, no other findings are necessary, and the appellate court,

\textsuperscript{126} \textit{U.S. v. Gaskell}, 985 F.2d 1056 (11th Cir. 1993).
\textsuperscript{127} \textit{Daskarolis v. Firestone Tire and Rubber Co.}, 651 F.2d 937, 940 (4th Cir. 1981).
\textsuperscript{128} \textit{Finchum v. Ford Motor Co.}, 57 F.3d 526, 530 (7th Cir. 1995).
\textsuperscript{129} \textit{Mittledier v. Chicago & N. W. Ry. Co.}, 441 F.2d 52, 55 (8th Cir. 1971).
\textsuperscript{130} \textit{Shipp v. General Motors Corp.}, \textit{supra}, n. 94.
applying an abuse of discretion standard of review, upholds the trial court’s ruling without further scrutiny. Because the standard is not based on any of the Federal Rules, litigants and courts are left in a quandry deciding how or whether to apply other provisions of the Federal Rules that ought to be considered.

By contrast, under Rules 403, 611, and 105, the court can—and should—also consider how much of the evidence to admit, which portions to exclude, and whether the potentially misleading effect of dissimilarities can be alleviated by cautionary instructions. These rules enable the court to consider the degree to which the test is likely to produce such a strong visceral response in jurors that it could overwhelm their objectivity, or whether the dissimilarities are explainable.

For example, in *Harvey v. General Motors Corp.*, the court held that it was proper to admit the test videos, but gave a cautionary instruction that the videos did not purport to “re-create” the accident and that the jury should be cognizant of the differences between the conditions of the test and those of the accident.132

131 *Muth, supra.*
132 873 F.2d 1343 (10th Cir.1989). The trial court gave the following cautionary instruction before permitting defendant to show the test films to the jury:

THE COURT: Members of the jury, there will now be displayed to you Defendant's Exhibit DD-2. In viewing this film, the Court has admitted it because it thinks that it would be helpful to you in understanding the oral testimony of Mr. Orlowski as well as the general principles of vehicle dynamics and occupant kinematics in patterns of injury mechanics to which Mr. Orlowski has testified.

But let me point out to you and instruct you that this involves a Chevrolet Impala [sic] with a solid roof, not a 1978 T-top Corvette, and you are not to ignore the distinctions between this demonstrative evidence and the actual event that is the subject matter of this action. You must make allowances for the differences between the actual event and the demonstrative evidence.”

*Id.*, at 1355.

See also, *Champeau v. Fruehauf Corp.* 814 F.2d 1271, 1278 (8th Cir.1987) (upholding admission of videotape of experiment with stipulation identifying the similarities and differences between experimental conditions and those involved in the accident); Mark Dombroff, *Innovative Developments in Demonstrative Evidence Techniques and Associated Problems of Admissibility*, 45 J. AIR LAW & COMM. 139, 146 (1979), discussion a cautionary instruction used in admitting a model of an airport in a trial arising from an air crash disaster.
In *Muth*, the Fifth Circuit upheld the discretion of the trial court without any explicit analysis of any of the applicable rules of evidence. With one minor exception, the Fifth Circuit’s entire discussion of the inadmissibility of the Malibu and CRIS tests failed even to cite any Federal Rule of Evidence.\(^{133}\) *Muth* is not an isolated case in this regard; numerous other federal cases have failed to cite a single Federal Rule of Evidence when applying the substantial similarity standard.\(^{134}\)

2. Vagueness of “Substantial Similarity.”

Given that it is impossible to conduct a test that is identical to the accident, how do courts make reasoned and logically consistent judgments about how similar a test needs to be in order to be “substantially similar”? The terminology of substantial similarity is so vague as to permit disregard of significant dissimilarities or to rationalize exclusion of tests based on trivial differences.\(^{135}\) How else to account for the fact that two courts reached precisely the opposite conclusion on the admissibility of the same test in the same kind of case? The same Malibu test video, excluded in *Muth*, was admitted in *Harvey*,\(^ {136}\) even though both cases presented rollover product liability claims based on allegations of inadequate roof strength. The imprecision of the

---

\(^{133}\) The sole reference was a footnote citation to a quotation from *Shipp v. General Motors Corp.*, 750 F.2d 418, 428 (5th Cir. 1985), which offers the conclusory opinion: “So viewed and so weighed on the Rule 403 scale, there was no error.” *Muth*, 461 F.3d, n.18 at 566.

\(^{134}\) See, e.g., Gladhill, supra; Chase, supra; McKnight, supra; Brandt, supra. But see, Jodoin, supra, 284 F.3d at 278.

\(^{135}\) Compare, *Randall v. Warnaco, Inc.*, 677 F.2d 1226, 1233-34 (8th Cir. 1982) (“A court may properly admit experimental evidence if the tests were conducted under conditions substantially similar to the actual conditions. Admissibility, however, does not depend on perfect identity between actual and experimental conditions. *Ordinarily, dissimilarities affect the weight of the evidence, not its admissibility.*”) (emphasis supplied, citations omitted), to *Gilbert v. Cosco Inc.*, 989 F.2d 399, 402 (10th Cir. 1993) (“[*E]xperiments which purport to recreate an accident must be conducted under conditions similar to that accident, while experiments which demonstrate general principles used in forming an expert's opinion are not required to adhere strictly to the conditions of the accident.”). (Emphasis supplied). See also, Comment, TRUTH, LIES AND VIDEOTAPE: ARE CURRENT FEDERAL RULES ADEQUATE?, 21 S.W. L. REV. 1199, 1210 (1992) (advocating that videotaped experiments should be admitted only when performed under conditions “substantially identical” to the incident in question); but see *Jodoin v. Toyota Motor Corp.*, 284 F.3d 272, 279 (1st Cir. 2002) (reversing trial court which required that vehicle in test be “virtually identical” with accident vehicle).
term, “substantially similar,” is compounded by the broad discretion appellate courts afford the trial court in making its determination.\textsuperscript{137}

No doubt, similarities and dissimilarities between a test and the actual conditions are a factor to be considered by a court, in determining: a) the reliability of the expert’s reasoning and methodology; b) the “fit” between an expert’s opinion and the facts of the case; c) the relevance of the test to the case; and d) whether the test is unfairly prejudicial or misleading under Rule 403. But even though dissimilarities are a part of the analysis, no one can seriously argue that they comprise the entire analysis.

A better analytical framework would include: 1) identification of the particular ways in which the test is similar and dissimilar to the events at issue; 2) the materiality of the dissimilarities;\textsuperscript{138} 3) whether the similarities are significant enough to render the test unreliable or not probative of any issue in dispute; 4) whether any of the dissimilarities are unfairly prejudicial and incapable of explanation or mitigation via cautionary instructions and cross examination; and 5) whether the test appears to be the “best evidence,” i.e., the availability (or lack thereof) of alternate test methodologies or protocols more reliable, more probative, and less likely to mislead the finder of fact; 6) whether the jury is more likely to be misled by the test, with its dissimilarities, than by evidence devoid of any testing at all; and 7) whether the party or expert is attempting to oversell the test as proving more than it actually does. It is noteworthy that most, if not all, of these considerations fall within the parameters of Rules 401, 403, 702 and 703.

3. False Dichotomy Between “Re-Creations” and “Abstract Principles.”

\textsuperscript{136} Harvey v. General Motors Corp., 873 F.2d 1343 (10\textsuperscript{th} Cir. 1989).
\textsuperscript{137} United States v. Rackley, 742 F.2\textsuperscript{nd} 1266, 1272 (11\textsuperscript{th} Cir. 1984); Wagner v. Int’l Harvester Co., 611 F. 2\textsuperscript{nd} 224, 232 (8\textsuperscript{th} Cir. 1979).
\textsuperscript{138} See Jodoin v. Toyota Motor Corp. 284 F.3d 272, 280 (1st Cir. 2002), which gives examples from prior cases to illustrate its conclusion that “when the relevant elements are sufficiently similar,” other dissimilarities can be
The substantial similarity standard creates a false, black-and-white, dichotomy between tests offered to “re-create” the accident in question and those offered to show an abstract principle. Frequently, however, there is a continuum rather than a bright line between the two. In *Muth*, one could view the Malibu and CRIS tests as both illustrating an abstract proposition and re-creating the accident, depending on how one poses the question. After all, these two tests tend to show both. Scientists devise tests which shed light on abstract and case-specific questions, as appears to have been the case with the Malibu and CRIS tests in *Muth*. The legal inquiry should not turn on whether the admissibility prize is to be found behind Curtain A or Curtain B. The Federal Rules make no distinction between abstract and case-specific tests. Rather, they require the court to assess the reliability and relevance of the test and to weigh the probative value against the prejudicial effect.

The substantial similarity standard, unlike the Federal Rules, provides no guideposts for resolving a situation in which a test would be admissible to prove an abstract proposition but inadmissible as a re-creation. By contrast, Federal Rule 105 provides that evidence which is admissible for one purpose and inadmissible for another may be admitted with appropriate instructions as to its limited purpose.\(^\text{139}\)

The permissive language of Rule 105, read in conjunction with Rule 403, suggests that such evidence should be admitted for a limited purpose, with appropriate limiting instructions, unless the court determines that its unfairly prejudicial impact could not be alleviated by instructions and that the prejudice substantially outweighs

\(^{139}\text{Rule 105 states:}

**Limited Admissibility**

When evidence which is admissible as to one party or for one purpose but not admissible as to another party or for another purpose is admitted, the court, upon request, shall restrict the evidence to its proper scope and instruct the jury accordingly.
whatever probative value the evidence might have. That sensible flexibility is unduly obscured by the all-or-nothing substantial similarity standard.

4. Exclusion of Demonstrations that are “Too Similar.”

If an expert conducts a test in a case in which there are disputed facts, whose version of the facts should the expert be permitted to rely upon when establishing the test parameters? For example, if defense witnesses claim that the plaintiff’s vehicle was traveling at 60 miles per hour, and plaintiff’s witnesses claim she was traveling 30 miles per hour, then at what speed should the expert run the test? *Muth* held that the trial court had discretion to deem Ford’s CRIS test as “prejudicial” because of the test’s similarities with Ford’s theory of the accident.\(^{140}\) In other words, because the test might be perceived as a re-creation in *some* respect, it was inadmissibly misleading because it was “too similar.”\(^{141}\) Such reasoning proves too much. It leads to a nonsensical analysis because, under its logic, any demonstration of any abstract principle is subject to the same criticism as long as the abstract principle is close enough to the facts of the disputed accident to be relevant at all. That the expert tried to conduct the CRIS test closer to the evidence made the test more probative, not more prejudicial, than if the test had been conducted in a manner wholly unrelated to any version of the facts.

Such logic, if followed to its logical end, leads to the anomalous conclusion that a test offered to demonstrate an abstract proposition is potentially inadmissible if it *is* substantially similar in *any* respect. Thus, if a plaintiff was injured by 20-pound object falling from a bridge and her expert sought to demonstrate Galileo’s test of dropping objects of different weights from the Tower of Pisa, the court could arguably exclude such a demonstration by arguing that one of

\(^{140}\) *Id.*

\(^{141}\) *Id.*
the objects weighed nearly the same as the object in the actual accident, or that they were
dropped from a similar height, or that they were shaped similarly, or that they were the same
color or that they were both made of rock.

5. Shifting of Foundational Burden.

The substantial similarity standard confuses the question of which party should bear the
foundational burden with respect to particular evidentiary concerns. Under the Federal Rules, the
proponent of expert testimony has the burden of showing that the evidence is relevant and
reliable under Rule 702.142 Once the proponent satisfies that burden, the objecting party has the
burden of showing that the probative value of the evidence is “substantially outweighed” by
various prejudicial factors.143

By contrast, the substantial similarity standard, as articulated by the Tenth Circuit in
Jackson v. Fletcher144 and generally followed by federal courts since then, places the entire
burden on the proponent. The Tenth Circuit's allocation of the burden did not result from any
analysis of any provision in the Federal Rules of Evidence. This was a serious oversight, given
that the Federal Rules, contrary to the common law rules for experimental evidence, created a
low threshold for relevance and shifted the burden to the opponent to show unfair prejudice in
order to exclude evidence that was otherwise reliable.

In some other situations, parties have argued that a common-law rule of exclusion should
overcome the Federal Rules’ more liberal rules of admissibility. For example, courts have
reached differing conclusions as to whether the common-law rule excluding evidence of a

---

143 U.S. v. Tse, 375 F.3d 148, 164 (1st Cir. 2004).
144 647 F.2d 1020, 1027 (10th Cir. 1981).
plaintiff’s collateral sources trumps the trial court’s discretion to admit such evidence under Rule 403. In a few other situations, the Federal Rules themselves explicitly shift the burden to the proponent to show the absence of unfair prejudice. Unfortunately, the courts’ reflexive application of the substantial similarity standard has pretermitted any analysis of whether the Rules’ silence with respect to experimental evidence should somehow negate the opponent’s burden of showing unfair prejudice under Rule 403.

The Tenth Circuit in *Jackson* decided to place the foundational burden entirely on the proponent simply by citing an older Tenth Circuit case, *Navajo Freight Lines v. Mahaffy*, which stated, “A party offering evidence of out of court experiments must lay a proper foundation by showing a similarity of circumstances and conditions.” Neither *Jackson* nor *Navajo Freight Lines* considered whether its analysis was consistent with the Federal Rules. *Navajo Freight Lines*, in turn, cited old cases that predated modern federal discovery and the Federal Rules of Evidence. Understandably, these old cases made no distinction between the burden of showing relevance and reliability and the burden of showing unfair prejudice. Indeed, the only case cited in *Navajo Freight Lines* that squarely addressed the issue, a 1906 case from Florida state court, said that evidence of experiments “should be received with caution, and only be admitted when it is obvious to the court, from the nature of the experiments, that the jury will be enlightened, rather than confused.” It held that “evidence of an experiment whereby to test the truth of testimony that a certain thing occurred is not admissible, where the conditions attending the

---

145 Compare e.g., *Sheehy v. Southern Pacific Transportation Co., Inc.*, 631 F.2d 649, 652 (9th Cir. 1980), with *Savoie v. Otto Candies, Inc.*, 692 F.2d 363, n. 8 at 371 (5th Cir. 1982) (collateral source rule).
147 174 F.2d 305, 310 (10th Cir. 1949).
alleged occurrence and the experiment are not shown to be similar. The similarity of circumstances and conditions go to the admissibility of the evidence, and must be determined by the court."\textsuperscript{149}

Rule 403, unlike the 1906 Florida case relied on by the Tenth Circuit, shifts the burden to the opponent. It authorizes exclusion only if the probative value of the evidence is “substantially outweighed” by its prejudicial effect and requires the objecting party to carry the burden of showing prejudice.\textsuperscript{150} The courts have given no reason why the burden of showing prejudice should be different for testing evidence than for other kinds of evidence.

In short, \textit{Muth} exemplifies the shortcomings of the substantial similarity standard and demonstrates why the First Circuit in \textit{Fusco} was right is describing the existing state of the law as “muddled.”\textsuperscript{151} The trial court in \textit{Muth} rejected the tests Ford offered as “not quite similar enough” yet “too similar.”\textsuperscript{152} The court expressed concern that the CRIS test would be confused with the subject accident because the test vehicle’s rotational speed and principal point of impact were consistent with the opinions of Ford’s experts, not plaintiff’s experts.\textsuperscript{153} According to the Fifth Circuit, this arguably made the CRIS test “too similar” to the accident facts. But if this was the court’s concern, then how could the court justify also excluding the video of the Malibu test, which was conducted using a completely different model vehicle, and using a test device that imposed no forward velocity at all? The court might have reasoned that the Malibu test was “not

\begin{footnotes}
\begin{itemize}
\item \textsuperscript{148} \textit{Hisler v. State}, 52 Fla. 30, 38, 42 So. 692, 695 (Fla.1906).
\item \textsuperscript{149} \textit{Id. Navajo Freight Lines} also cited a 1936 California state court of appeal decision which held that “it must always be first shown whether circumstances are similar or so similar as to render result of experiments illustrative of the questions under consideration,” but that case did not discuss whether the proponent's burden also extended to questions relating to the prejudicial impact of such experiments. \textit{Collins v. Graves}, 17 Cal.App.2d 288, 299, 61 P.2d 1198, 1203 (Cal.App.1936).
\item \textsuperscript{150} \textit{Tse, supra}, n. 143.
\item \textsuperscript{151} \textit{Fusco v. General Motors Corp., supra}, 11 F.3d 259 at 263-264.
\item \textsuperscript{152} \textit{Muth}, at 566.
\end{itemize}
\end{footnotes}
similar enough,” although it did not say this, but the court did not analyze the admissibility of the two tests separately. If the court’s holding, that the tests were “too similar yet not similar enough,” sounds like the porridge tasted by Goldilocks, it is. The substantial similarity standard employed by the court makes it impossible for a litigant to determine what test, if any, is “just right.”

Finally, the court in *Muth* distinguished between the admissibility of the tests themselves and photographs and video of those tests. The trial court excluded the latter but not the former. No one contested that the video and photographs helped the jury understand the general dynamics involved in rollover accidents; that they illustrated Ford’s claim that, during rollovers, head and neck injuries can occur prior to roof deformation. It is unclear from the opinion whether plaintiff ever objected to the tests as unreliable. However, one must presume that the trial court probably would have excluded evidence of the tests altogether had it found the tests to be unreliable or irrelevant. But the trial court apparently made no such determination. Rather, it permitted Ford’s expert to testify about the tests but not to see any of what the court characterized as “his visual aids.”

No doubt, there may be circumstances under which the court should admit evidence of the test but exclude the demonstrative evidence. But such a decision ought to be explicitly based upon application of Rules 403, 611(a), 703 and 105 rather than the substantial similarity standard.

**F. The Solution: Follow the Rules.**

As previously suggested, the Federal Rules already provide the necessary criteria for

---

153 *Id.*, at 567.

154 *Muth*, at 566.
assessing the admissibility of experimental evidence. They offer a more detailed and systematic methodology for identification and analysis of the pertinent issues pertaining to such evidence. 

*Muth* presents an excellent example of the reasons that the substantial similarity standard can distract courts and litigants from analyzing the issues appropriately. It also highlights the need for making clear that the Federal Rules can and should apply to demonstrative evidence, just as with other kinds of evidence. Thus, as with other kinds of evidence, the methodology for admissibility of demonstrative evidence of testing should be as follows:

1) Has the proponent demonstrated that the tests were reliable and relevant under Rules 702 and 401? In the case of demonstrative evidence, “relevant” ought to be understood to mean that they fairly and accurately explain, illustrate, or clarify, other admissible evidence.156

2) If reliable and relevant, has the objecting party shown that the relevance is substantially outweighed by the dangers of unfair prejudice as set forth in Rule 403?

3) If the tests are dissimilar in some fashion from the accident, are the dissimilarities material? Can the effect of these dissimilarities be alleviated by cautionary instructions, or is the test likely to evoke a visceral response in jurors that cannot be effectively overcome by cautionary instructions and vigorous cross examination?

4) Is the demonstrative evidence being used to bring in evidence that is otherwise inadmissible under Rule 703, and, if so, has the proponent shown that the probative value of such evidence in assisting the jury to evaluate the expert's opinion substantially outweighs its prejudicial effect?

How a court reconciles the applicable provisions of the Federal Rules with a common-

155 *Id.*
156 Brain & Broderick, at 923.
law standard that predates the Federal Rules is well defined. Rule 402 of the Federal Rules declares that “[a]ll relevant evidence is admissible, except as otherwise provided by the Constitution of the United States, by Act of Congress, by these rules, or by other rules prescribed by the Supreme Court pursuant to statutory authority.”\footnote{157} It follows that common-law evidentiary decisions should not trump the Federal Rules.

\textit{Daubert} presents a prime example. The Court was asked to reconcile the pre-F.R.E. \textit{Frye} test with Rule 702. Some courts had continued to apply \textit{Frye} even after enactment of the Federal Rules. The Court noted that the legislatively-enacted Federal Rules should be interpreted in the same manner as any statute.\footnote{158} Citing \textit{United States v. Abel},\footnote{159} the Court concluded that “the Rules occupy the field,” but that “the common law nevertheless could serve as an aid to their application.”\footnote{160} The Court contrasted the situation in which the common-law rule was “entirely consistent” with the provision in the Federal Rules, and one in which the Court “was unable to find a particular common-law doctrine in the Rules.”\footnote{161} In the former situation, the body of law contained in the common-law decisions could continue to provide a source of guidance; in the latter, the common-law rule was “superseded.”\footnote{162} Because nothing in the Federal Rules incorporated the \textit{Frye} test and because there was no indication that the Federal Rules as a whole or any specific rule was intended to incorporate it, the assertion that the Federal Rules were intended to “assimilate” the \textit{Frye} test was “unconvincing.”\footnote{163} Accordingly, the Court abandoned the \textit{Frye} standard and, instead, established a new standard consistent with the language and

\footnote{157} \textit{U.S. v. Curtin}, 2007 WL 1500295, *7 (9th Cir.2007).
\footnote{159} 469 U. S. 45 (1984).
\footnote{160} 509 U.S., at 587.
\footnote{161} \textit{Id.}, at 588.
\footnote{162} \textit{Id.}
\footnote{163} \textit{Id.}
purposes of the Federal Rule.\textsuperscript{164}

The same analysis ought to lead to a similar conclusion as to the substantial similarity standard for experimental evidence. The substantial similarity standard is found nowhere in the Federal Rules. The Federal Rules provide no basis for continuing to treat experimental evidence as a discrete category of evidence, as it was before. The reliability, relevance, and prejudice provisions of the Federal Rules encompass the issues raised by such evidence. It follows that, as with \textit{Frye}, the pre-F.R.E. substantial similarity standard should give way to the standards already incorporated in the Federal Rules.

\textbf{CONCLUSION}

The substantial similarity standard for evidence of tests, like the \textit{Frye} standard, is an artifact of pre-Federal Rules evidence law. Treating such evidence as “experimental,” like \textit{Frye’s} treatment of “novel scientific” evidence, should no longer trigger special rules of admissibility. In a post-\textit{Kumho} world, the artifacts of common-law “experimental” evidence should not continue to confuse, let alone preempt, the relevance, reliability and prejudice criteria of the Federal Rules. No doubt the Rules would benefit from clarification of the standards for admissibility of demonstrative evidence, to bring it more clearly under the umbrella of the existing rules, such as was proposed by \textsc{Brain & Broderick}.

Dean Wigmore got it mostly right long ago, when he defined the issue as follows:

\begin{quote}
[\textit{W}hether data whose relevance is not questioned are objectionable or inferior because they have been obtained, not by observing merely such casual material as
\end{quote}

\textsuperscript{164} Another example of a common-law evidence rule supplanted by the Federal Rules of Evidence is the doctrine of \textit{res gestae}. \textit{F.D.I.C. v. Fidelity & Deposit Co. of Maryland}, 45 F.3d 969, 979 (5\textsuperscript{th} Cir. 1995). This “obsolete” doctrine was described by other federal courts as “useless, harmful, and almost inescapable of a definition.” \textit{Stephens v. Miller}, 13 F.3d 998, 1003 (7\textsuperscript{th} Cir.) (and cases cited therein) (en banc), \textit{cert. denied}, 513 U.S. 808 (2004).
nature has provided for us, but by carefully arranging the conditions so as to obtain by experiment trustworthy results. That there should be such a distinction between observation and experiment would be unworthy of our law.\textsuperscript{165}

Today’s Federal Rules no longer require an act of faith that the expert has “carefully arranged” conditions to produce “trustworthy results.” The rules provide ample opportunity, in advance of trial, to gauge the reliability of the test, to decide whether it “fits” the issues of the case, whether it is probative of a material issue, and whether its relevance is substantially outweighed by a variety of prejudicial factors. If these rules are properly applied, an amorphous substantial similarity standard is obsolete. Because it distracts from the standards for admissibility set forth in the Federal Rules, substantial similarity should no longer be employed as the litmus test for admissibility of experimental evidence.

\textsuperscript{165} II Wigmore on Evidence (Chadbourn Ed. 1979), § 445.