If the Glove Don't Fit,** Update the Glove: The Unplanned Obsolescence of the Substantial Similarity Standard for Experimental Evidence

Jonathan M. Hoffman
Martin Bischoff Templeton Langslet & Hoffman, LLP, in Portland, Oregon

Follow this and additional works at: http://digitalcommons.unl.edu/nlr

Recommended Citation
Jonathan M. Hoffman, If the Glove Don't Fit,** Update the Glove: The Unplanned Obsolescence of the Substantial Similarity Standard for Experimental Evidence, 86 Neb. L. Rev. (2007)
Available at: http://digitalcommons.unl.edu/nlr/vol86/iss3/3
Jonathan M. Hoffman*

If the Glove Don’t Fit,** Update the Glove: The Unplanned Obsolescence of the Substantial Similarity Standard for Experimental Evidence

TABLE OF CONTENTS

I. Introduction ........................................... 634
II. Discussion ............................................ 636
   A. The Muth Decision .................................... 636
      1. The Malibu and CRIS Tests ......................... 637
      2. The Fifth Circuit's Decision on Admissibility .... 640
   B. Origins of the Substantial Similarity Standard .... 641
      1. Pre-Federal Rules Cases ............................. 642
      2. Post-Federal Rules Cases ............................. 644
   C. Demonstrative Evidence—Whatever It Is ............. 647
   D. Experimental Evidence and the Federal Rules ........ 650
      1. Rule 702 ........................................... 650
      2. Rule 403 ........................................... 653
      3. Rule 703 ........................................... 656

© Copyright held by the NEBRASKA LAW REVIEW.

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 Hon. Robert E. Jones, Prof. Laird Kirkpatrick, and Derek Whitefield of Dykema Gossett LLP for their insights and critiques. Any errors in the analysis contained in this article are the author’s own.

** The phrase comes from Johnny Cochran’s closing argument in the O.J. Simpson murder trial. Cochran actually said, “if it doesn’t fit, you must acquit.” CNN.com, If it Doesn’t Fit, You Must Acquit, www.cnn.com/US/OJ/daily/9-27/8pm/ (last visited March 11, 2008). However, by the time of Cochran’s death in 2005, the statement in popular culture had morphed into the ungrammatical: “If the glove don’t fit, you must acquit.” See, e.g., The BoxHeads.net, If the Glove Don’t Fit, You Must Acquit, www.theboxheads.net/modules.php?name=News&file=print&sid=90 (last visited March 11, 2008). The widely-perceived failure of the prosecution’s in-court experiment may have played a role in Simpson’s acquittal.
I. INTRODUCTION

In the wake of Kumho Tire Co. Ltd. v. Carmichael, more forensic 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 is often crucial. The stakes are high. The cost of conducting proper testing can be considerable, but the cost of losing the case if a party fails to test (or if the court excludes the test) is enormous. In large measure, the jury’s verdict may rest on demonstrative evidence 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, so is the frequency with which the admissibility of evidence of testing is the decisive issue in the case. This is partly because of the effect of Kumho, but it is also due to the increasing sophistication of forensic testing and demonstrative evidence.

Although many of the leading cases arise in automotive product liability cases, the issue transcends that area of civil litigation and encompasses a broad range of cases. 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 U.S. Court of Appeals for the Fifth Circuit’s 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 to Rule 702 of the Federal Rules of Evidence (hereinafter “Federal Rules”), experts typically are required to test their hypotheses if pos-

2. 461 F.3d 557, 565–67 (5th Cir. 2006).
3. Rule 702 states:
   Testimony by Experts.
sible, to conduct the test in accordance with the rigors of their profession, and to apply the results reliably to the facts. But, if evidence of tests satisfying those requirements is offered at trial, courts continue to rule unpredictably on the admissibility of those tests.4

In civil litigation, unpredictability is an equal opportunity source of frustration for plaintiffs and defendants alike. The unsuccessful proponent of excluded evidence may be a plaintiff5 or a defendant.6 Likewise, the unsuccessful opponent of admitted evidence may be either a plaintiff or a defendant.7

There is one principal reason for the current state of unpredictability. Courts continue to treat experimental evidence as a discrete category of evidence, the admissibility of which is judged by a nebulous substantial similarity standard that predates the Federal Rules and no longer serves a useful purpose. Worse yet, the courts have extended the same substantial similarity standard to demonstrative evidence, with no serious consideration of the appropriateness of doing so.

Rigorous application of the Federal Rules would ensure that experimental evidence is reliable and relevant (as required by Rules 702 and 401) and that risks of unfair prejudice do not outweigh that evidence's relevance (as required by Rule 403).8 When courts respect the

---

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.

FED. R. EVID. 702.

4. See, e.g., Shipp v. Gen. Motors Corp., 750 F.2d 418 (5th Cir. 1985) (addressing the trial court's admission of plaintiff's evidence and exclusion of defendant's similar evidence). Also compare Muth v. Ford Motor Co., 461 F.3d 557 (5th Cir. 2006) with Harvey v. General Motors Corp., 873 F.2d 1343 (10th Cir. 1989) (reaching diametrically opposed conclusions on the admissibility of the same test in the same kind of products liability case).

5. See, e.g., Jodoin v. Toyota Motor Corp., 284 F.3d 272 (1st Cir. 2002); Finchum v. Ford Motor Co., 57 F.3d 526, 530 (7th Cir. 1995).


8. Rule 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.

FED. R. EVID. 403.
constraints that now exist under these, and other, Federal Rules, the
imprecise substantial similarity standard from the past becomes un-
necessary. Continued use of the substantial similarity standard in-
jects needless confusion and imprecision into evidentiary decision
making. Although similarities and dissimilarities between test condi-
tions and accident conditions are a pertinent factor in assessing the
reliability, relevance, and prejudicial impact of tests conducted by ex-
erts, the old substantial similarity standard for determining the ad-
missibility of such evidence distracts litigants and courts from more
refined 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.

Nor should the substantial similarity standard govern the admissi-
bility of demonstrative evidence. Demonstrative evidence is simply
not the same thing as experimental evidence. Rule 401 ought to be
understood to provide, or amended to make explicit, that “relevant evi-
dence” includes evidence that fairly and accurately explains, illus-
trates, or clarifies other admissible evidence. With such a
clarification, Rule 403 and the other Federal Rules provide a more
suitable framework for judging the admissibility of demonstrative evi-
dence that flows from testing than the old substantial similarity
standard.

An expert’s test should not be required to be “substantially similar”
to be admissible. Rather, it ought to be reliable within the meaning of
Rule 702 and be probative of the hypothesis it was intended to prove
or disprove. Nor should demonstrative evidence be treated under a
substantial similarity standard. Muth is not the first case to have
stumbled into this thicket. We would all be better off if it were the
last. It is time to give the substantial similarity standard a decent
burial.

II. DISCUSSION

A. The Muth Decision

Muth v. Ford Motor Co. demonstrates the shortcomings of the
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. The plaintiff
brought a product liability claim against Ford, contending that the vehi-
cle’s roof was insufficiently strong, and that the plaintiff was injured
as a result of the roof crush. Ford asserted that a stronger roof
would not have prevented head and neck injuries in rollover accidents

9. 461 F.3d 557 (5th Cir. 2006).
10. Id. at 560.
11. Id.
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.\textsuperscript{12} 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.\textsuperscript{13}

1. The Malibu and CRIS Tests

Although numerous scientific publications have addressed the role of roof strength in protecting occupants in rollover collisions, and despite years of test development, it has been extremely difficult to develop a repeatable dynamic rollover test method.\textsuperscript{14} This is a major reason why the National Highway and Traffic Administration’s (“NHTSA”) regulatory standard continues to specify a static, rather than dynamic, test for roof strength.\textsuperscript{15} Repeatability is difficult because a motor vehicle rolling over is somewhat analogous to tossing a football onto the ground—no one can predict with certainty where it will bounce. The Malibu and Controlled Rollover Impact System (“CRIS”) tests are two of the foremost dynamic evaluations of occupant kinematics in rollover accidents.

General Motors designed the Malibu tests in the early 1980s.\textsuperscript{16} GM conducted two series of full-scale rollover crash tests at the General Motors Proving Grounds. The first series used unrestrained instrumented anthropomorphic dummies in 1983 Chevrolet Malibus.\textsuperscript{17} The primary expressed purpose was “to investigate the conditions leading to injuries in rollover conditions and the effect of roof strength

\begin{itemize}
\item \textsuperscript{12} Id. at 560–61.
\item \textsuperscript{13} Id. at 561.
\item \textsuperscript{16} Muth, 461 F.3d at 565.
\item \textsuperscript{17} Kenneth F. Orlowski, R. Thomas Bundorf & Edward A. Moffatt, Rollover Crash Tests—The Influence of Roof Strength on Injury Mechanics, in SAE TECHNICAL PAPER SERIES, at 182 (SAE Technical Paper Series No. 851734, 1985) [hereinafter “Malibu I”]. These tests are also discussed in Muth, 461 F.3d at 561 and Harvey v. Gen. Motors Corp., 873 F.2d 1343, 1355 (10th Cir. 1989).
\end{itemize}
in preventing injuries."¹⁸ GM subsequently ran a second series of tests using belted dummies.¹⁹ In both series, test films depicted 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 m.p.h. The 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.²⁰

Figure 1, below, illustrates the layout of the test.

![Figure 1: Malibu Tests](image)

(Figure 1 is reprinted with permission from SAE Paper 902314 ©1990 SAE International)

The tests measured loads imposed on instrumented dummies and compared the results between production vehicles and vehicles modified to make the roof rigid.²¹ As illustrated by Figure 2, below, in both series of tests, the peak neck load occurred at the beginning of the collision before any significant roof deformation had occurred.²²

---

¹⁸. Malibu I, supra note 17, at 185.
²¹. Malibu I, supra note 17, at 182; Malibu II, supra note 19, at 103.
²². Malibu I, supra note 17, at 192; Malibu II, supra note 19, at 106; see also Michael B. James, Ronald P. Nordhagen, Dennis C. Schneider, Sung-Woo Koh, Occupant Injury in Rollover Crashes: A Re-Examination of Malibu II, in SAE TECHNICAL PAPER SERIES (SAE Technical Paper Series No. 2007-01-0369, 2007).
creasing the roof’s rigidity did not significantly reduce the dummies’ neck loads. Both series of tests led to the authors’ conclusion that roof crush simply reflects the magnitude and location of the impact of the roof with the ground. The harder the hit, the greater the deformation and the greater the potential for injury. Increasing the roof strength will reduce deformation but it will not reduce neck loading when the occupant is at the point of impact.

Figure 2: Timing of Peak Neck Load and Roof Crush

(Figure 2 is reprinted with permission from SAE Paper 902314 ©1990 SAE International)

The CRIS test, conducted by Exponent, Inc., during 2000–2001, was developed in response to the NHTSA’s 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.” The CRIS test released a rotating vehicle in a precise manner onto the ground from the back of a moving semi-trailer. “By controlling the roll, pitch and yaw angles, translation and vertical velocities, and roll velocity of the

23. Malibu I, supra note 17, at 192; Malibu II, supra note 19, at 106.
26. Muth, 461 F. 3d at 566.
vehicle, the first roof-to-ground interaction was repeatable from test to test.”

NHTSA did not adopt CRIS (or any other dynamic test protocol) when it last revised the federal safety standards in 2001 because the vehicle's motion after first contact with the ground was not sufficiently repeatable; however, the agency acknowledged that the CRIS test was “helpful in understanding occupant kinematics during rollover crashes.”

Figure 3: Schematic Diagram of the CRIS

(Figure 3 is reprinted with permission from SAE Paper 2003-01-0172 ©2003 SAE International)

Like the Malibu tests, the CRIS tests compared matched pairs of vehicles: 1998–2000 Ford Crown Victorias, one a production vehicle and the other equipped with a roll cage. As with the Malibu tests, the peak neck axial loads measured on instrumented dummies were nearly identical in the roll-caged and production vehicles. Moreover, as with the Malibu tests, the peak head accelerations and neck loads occurred before any significant roof deformation. The authors concluded that roof deformation “does not affect injury potential for belted occupants whose heads are at or near the roof at the time and location of roof-to-ground impact.”

2. The Fifth Circuit's Decision on Admissibility

In Muth, Ford offered both tests to show that spinal cord injuries in rollover accidents are not caused by roof crush and to assist the jury in understanding Ford's experts' testimony regarding the general dy-

29. Moffatt et al., supra note 14.
30. Id.
31. Id.
32. Id.
namic characteristics of rollover accidents.\textsuperscript{33} The trial court permitted the experts to testify but excluded demonstrative evidence of the tests, including film and photographs. It excluded the test films because it concluded that the tests were not conducted under substantially the same conditions as those involved in the lawsuit.\textsuperscript{34}

Ford appealed a judgment for the plaintiff, but the Fifth Circuit affirmed. Writing for the court, Judge Higginbotham recited the rule that when "demonstrative evidence is offered only as an illustration of general scientific principles,\textsuperscript{35} not as a re-enactment of disputed events, it need not pass the substantial similarity standard."\textsuperscript{36} The circuit court cautioned, however, that such demonstrative evidence must not be misleading in and of itself.\textsuperscript{37} A demonstration might mislead when it resembles the disputed accident, but, the court observed, this resemblance gives rise to the requirement of substantial similarity.\textsuperscript{38} 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.\textsuperscript{39}

Ford claimed that the CRIS test was not offered as a re-enactment but only to show general scientific principles.\textsuperscript{40} 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 reconstruction evidence offered by plaintiffs.\textsuperscript{41} 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."\textsuperscript{42} Accordingly, the Fifth Circuit affirmed the trial court's exclusion of the test videos and photographs.

B. Origins of the Substantial Similarity Standard

The Fifth Circuit opinion in \textit{Muth} provides no indication that the parties or the court ever questioned the rationale of applying a sub-

\begin{itemize}
\item \textsuperscript{33} 461 F.3d 557, 561 (5th Cir. 2006).
\item \textsuperscript{34} \textit{Id}.
\item \textsuperscript{35} \textit{See} \textit{Gilbert v. Cosco}, Inc., 989 F.2d 399, 401-02 (10th Cir. 1993); \textit{Four Corners Helicopters, Inc. v. Turbomeca, S.A.}, 979 F. 2d 1434, 1441 (10th Cir. 1992); \textit{Swajian v. Gen. Motors Corp.}, 916 F.2d 31, 36 (1st Cir. 1990) (disallowing admission of videotaped tests because defendant "fail[ed] to mention even one such general scientific principle that the videotapes purport[ed] to demonstrate").
\item \textsuperscript{36} \textit{Muth}, 461 F.3d at 566.
\item \textsuperscript{37} \textit{Id}.
\item \textsuperscript{38} \textit{Id}.
\item \textsuperscript{39} \textit{Id}. at 566-67.
\item \textsuperscript{40} \textit{Id}.
\item \textsuperscript{41} \textit{Muth v. Ford Motor Co.}, 461 F.3d 557, 567 (5th Cir. 2006).
\item \textsuperscript{42} \textit{Id}.
\end{itemize}
substantial similarity standard to determine the admissibility of demonstrative evidence of tests that "re-enact" the accident. However, such a standard is found nowhere in the Federal Rules or their comments. It predates both Daubert\textsuperscript{43} and the Federal Rules of Evidence. Not surprisingly, the rule's origins are unrelated to existing standards for admissibility of expert testimony. Rather, a century before the codification of the Federal Rules, common law decisions addressed the admissibility of "experimental evidence."\textsuperscript{44} Thereafter, as seen below, courts began to treat experimental evidence as a discrete category and used a substantial similarity standard to determine its admissibility. Moreover, the standard governed the admissibility of evidence of the experiment itself rather than demonstrative evidence illustrating or depicting the experiment.

1. Pre-Federal Rules Cases

In one early example of "experimental evidence," May Department Stores Co. v. Runge,\textsuperscript{45} the plaintiff sustained an on-the-job injury when a 200-pound "truck" fell down an elevator shaft and landed on him.\textsuperscript{46} The principal issue at trial was whether the truck had somehow passed through a gate on an upper floor of the building.\textsuperscript{47} The parties disputed whether there had been sufficient clearance below the gate for the truck to pass through.\textsuperscript{48} The gate was raised and lowered by a rope.\textsuperscript{49} 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.\textsuperscript{50}

The defendant disputed this, claiming that the truck was probably in the elevator shaft when the plaintiff used the elevator.\textsuperscript{51} It offered the testimony of fact witnesses who, after the accident, attempted to raise the gate as high as it could go.\textsuperscript{52} They tied the rope so that the gate would be raised as high as possible, consistent with the automatic operation of the elevator.\textsuperscript{53} Then they tried to push a truck of the same dimensions as the truck that injured the plaintiff under the

\textsuperscript{44} Darling v. Westmoreland, 52 N.H. 401 (1872).
\textsuperscript{45} 241 F. 575 (8th Cir. 1917).
\textsuperscript{46} Id. at 576–77.
\textsuperscript{47} Id. at 577.
\textsuperscript{48} Id.
\textsuperscript{49} Id.
\textsuperscript{50} May Dep't Stores Co. v. Runge, 241 F. 575, 577 (8th Cir. 1917).
\textsuperscript{51} Id.
\textsuperscript{52} Id.
\textsuperscript{53} Id.
gate into the elevator shaft. They found that it was impossible to do so. The trial court rejected defendant’s evidence as “self-serving,” but the Eighth Circuit Court of Appeals reversed. The appellate court concluded that the evidence was no more self-serving than other evidence offered by the plaintiff and that it was admissible as experimental evidence, stating:

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 casualty looked at the articles in some past time.

Citing the 1872 decision of Darling v. Westmoreland, the court concluded that such evidence was admissible if it “substantially tends to establish the fact it is offered to prove.” May Stores cited older cases 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.” 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.

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. Nor was there any attempt to present the experiment to the jury via demonstrative evidence.

The court in May Stores did not explicitly adopt a “substantial similarity” standard for the admissibility of the experiment. Its requirement that the conditions of the tests be the same as those of the subject accident “as near as practicable” is susceptible to a different interpretation, based on the practicality of achieving greater similarity rather than the extent of the dissimilarities. Moreover, May Stores treated the relevancy of the experiment as an integral element of the admissibility determination, given the court’s statements that the evidence must be probative (“substantially tends to establish the fact it is offered to prove”) and not unfairly prejudicial (“where it is likely to

54. Id. at 577-78.
55. May Dep’t Stores Co. v. Runge, 241 F. 575, 578 (8th Cir. 1917).
56. Id.
57. Id. at 578.
58. Id. at 578-79 (quoting Darling v. Westmoreland, 52 N.H. 405 (N.H. 1872)).
59. Id. at 579.
61. May, 241 F. at 578.
62. Id. at 579.
tend more to confusion and inconvenience than to justice and certainty”). However, other cases predating *May Stores* applied a standard more akin to substantial similarity. In *Tackman v. Brotherhood of American Yeomen*, for example, the Court concluded:

"[I]t is now well settled that, when the conditions are shown to be substantially the same, evidence of actual experiment is an acceptable aid in determining the issues in a case. Indeed, there is an advantage in experiments over some other kinds of evidence, in that it [sic] substitutes experience for speculation and a demonstration in place of what sometimes seems no more than guesswork. . . [.but] the conditions must be such that they may be found to have been not only possible, but reasonably probable . . . ."  

Courts continued to apply a substantial similarity standard for the admission of experimental evidence thereafter without focusing on whether the evidence was presented through fact witnesses or experts. More recently, but still decades before adoption of the Federal Rules, the Third Circuit Court of Appeals 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.  

In each of the above cases, courts applied the standard of evidence to the tests themselves, as contrasted with the Fifth Circuit’s application of the standard in *Muth* solely to demonstrative evidence of the tests. Although other pre-Federal Rules cases applied the substantial similarity standard to in-court experiments or demonstrations, using the same rationale as that used for out-of-court tests, there is little if any analysis whether the same standard ought to apply to both. Nor could any analysis be found, prior to *Muth*, for permitting oral testimony of the tests relied upon by the experts but excluding demonstrative evidence of those tests. Indeed, if a substantial similarity test governed both, one would think the court would exclude all evidence of the tests or let it all in.

2. Post-Federal Rules Cases

Following adoption of the Federal Rules, some courts continued to treat “experimental evidence” as a discrete category of evidence, without questioning whether such treatment was consistent with the Federal Rules. The Tenth Circuit Court of Appeals’ decision in *Harvey*

63. *Id.*
64. 106 N.W. 350 (Iowa 1906).
65. *Id.* at 351.
69. See, e.g., McKnight v. Johnson Controls, Inc., 36 F.3d 1396, 1402 (8th Cir. 1994).
v. General Motors,70 for example, addressed the admissibility of the same Malibu tests later at issue in Muth, by continuing to apply the pre-Federal Rules "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.71

In Fusco v. General Motors Corp.,72 the plaintiff claimed that the vehicle lost control and struck a telephone pole because of fatigue failure of a ball stud in the front suspension.73 GM countered that the ball stud failure was a result, not a cause, of the collision with the pole.74 GM's 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."75 The trial judge excluded the test, and the First Circuit Court of Appeals affirmed.76 It characterized the issue as "interesting and important" and acknowledged that the case law surrounding such evidence was "muddled."77 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, 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.78

The First Circuit characterized this rule as a doctrine, "predating and now loosely appended to Rule 403."79 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,

70. 873 F.2d 1343 (10th Cir. 1989).
71. Id. at 1356 (quoting Jackson v. Fletcher, 647 F.2d 1020 (10th Cir. 1981)). Curiously, the court in Muth did not cite Harvey, even though it addressed the admissibility of one of the same tests excluded in Muth.
72. 11 F.3d 259 (1st Cir. 1993), cited in Muth v. Ford Motor Co., 461 F.3d 557, 566 n.20 (5th Cir. 2006).
73. Id. at 260-61.
74. Id. at 261.
75. Id.
76. Id. at 264.
78. Id. at 263-64.
79. Id. at 264. See the text of Rule 403 supra note 8.
even inconsistencies, between the two.\textsuperscript{80} The First Circuit's opinion also attempted to link the substantial similarity standard to the Supreme Court's decision in \textit{Daubert} because Rule 702 imposed a reliability requirement for expert testimony.\textsuperscript{81} Although not expressly articulated by the court, the implication was that a re-creation of an accident that was not substantially similar could be excluded as unreliable.

Historically, many in-court "experiments" were far less elaborate and scientific than those at issue in \textit{Fusco}, \textit{Harvey}, and \textit{Muth} and did not necessarily involve experts. In fact, some did not even involve witnesses but, rather, were simply performed by the lawyers and 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."\textsuperscript{82} 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.\textsuperscript{83}

Many of the reported cases arose from attempts by parties to "re-create" the accident or incident at issue in the case, with varying degrees of verisimilitude. The courts treated accident reconstructions or re-creations with suspicion, and expert testimony to reconstruct or recreate an accident was itself viewed with substantial skepticism, if not outright mistrust. Such evidence was variously considered conjectural, speculative, and/or misleading, and judges feared it could invade the province of the jury in resolving disputed facts.\textsuperscript{84} 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. For example, such evidence may be used to challenge the non-physical abilities of witnesses (e.g., to be able to see, hear, or smell)\textsuperscript{85} or to

\textsuperscript{80} See discussion, infra, Section II.D.2.
\textsuperscript{81} \textit{Fusco}, 11 F.3d at 264.
\textsuperscript{85} An illustrative example from popular culture is depicted in the classic film, \textit{THE YOUNG PHILADELPHIANS} (Warner Bros. 1959), in which the ambitious young tax lawyer, portrayed by a youthful Paul Newman, turns into a criminal defense law-
challenge a purportedly common-sense proposition.\textsuperscript{86} 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.\textsuperscript{87} However, there is a dearth of decisions in which courts have directly considered whether the pre-Federal Rules treatment of experimental evidence is appropriate, necessary, or consistent with the Federal Rules. In short, courts have usually relied on \textit{stare decisis} rather than on critical analysis as to whether the case law is consistent with the Federal Rules.

C. Demonstrative Evidence—Whatever It Is

In \textit{Muth}, the trial court excluded demonstrative evidence of the Malibu and CRIS tests but permitted expert testimony about the tests. In effect, the court imposed a more stringent admissibility threshold for demonstrative evidence than for other evidence. It is unclear why this should be the case. Indeed, there was no explicit discussion concerning whether the substantial similarity test should apply at all to demonstrative evidence.

Absent any standards for the admissibility of demonstrative evidence, courts lack clear criteria for determining the proper treatment of such evidence. Indeed, there is some confusion whether the admissibility threshold for demonstrative evidence ought to be higher or lower than that for the underlying proposition. For example, although the court in \textit{Muth} implicitly set a higher standard for demonstrative evidence, 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; it therefore is not required to meet the higher post-\textit{Daubert} admissibility standards.\textsuperscript{88}

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 contain neither a definition of demonstrative evidence nor any rule spelling out how a trial court should determine its admissibility or


\textsuperscript{87} See Champeau v. Fruehauf Corp., 814 F.2d 1271, 1278 (8th Cir. 1987) (discussing rules for admission of experimental evidence); see also, Jodoin v. Toyota Motor Corp., 284 F.3d 272 (1st Cir. 2002); Chase v. Gen. Motors Corp., 856 F.2d 17 (4th Cir. 1988); McKnight v. Johnson Controls, Inc., 36 F.3d 1396 (8th Cir. 1994); Gladhill v. Gen. Motors Corp., 743 F.2d 1049 (4th Cir. 1984).

other use at trial. Historically, demonstrative evidence was "not evidence" and the rules did not apply as long as it was not given to the jury during deliberations. The only mention of the subject in the Federal Rules consists of 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 additional confusion because of the absence of a cross-reference to any other rule that might bear upon the use of demonstrative evidence.

Most of the leading publications on the subject of demonstrative evidence were authored by prominent practitioners, most notably Melvin Belli, rather than legal scholars, and a "concomitant avalanche" of others as well. Such writings popularized demonstrative evidence rather than analyzed it, and failed to present any "unifying theory explaining its relevance, admissibility, or use." One book seriously discussed the evidentiary rules with respect to a wide variety of types

---


91. Id. at 961 n.7, 1017 n. 206. 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.

92. 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 harassment or undue embarrassment.

FED. R. EVID. 611(a). See also United States v. Johnson, 54 F.3d 1150, 1158 (4th Cir. 1995).

93. The Advisory Committee Notes to Rule 611(a) 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)."

94. See, e.g., MELVIN BELLI, MODERN TRIALS (1954); READY FOR THE PLAINTIFF! A STORY OF PERSONAL INJURY LAW (1956).

95. See Brain & Broderick, supra note 89, at 999–1000 n.46 (listing sources).

96. Id. at 1002.
of demonstrative evidence.97 But this book said little more than that
the admission of demonstrative evidence is discretionary and that the
use of such evidence is limited by Rule 403. This is not a criticism of
the book but an indication of the lack of clear guidance either in the
Federal Rules or case law.

Scholarly writings are of little additional help. Wright & Graham's
treatise merely suggests that the best that can be done is to recognize
that demonstrative evidence "is both within and without the province
of evidence" and proposes that the courts apply the Federal Rules
"when they supply a workable solution to the problem at hand" and
"resort to other procedural regulations" when the Federal Rules are
"inadequate to cope with the question."98 Unfortunately, it provides
no suggestions for determining when or how the Federal Rules "sup-
ply a workable solution" to a type of evidence they ignore, nor does it
indicate what "procedural regulations" would fill such gaps.

Brain and Broderick's excellent article highlights many of the
problems with the publications of their forebears. They note the ab-
ence of any scholarly theory explaining the relevance of demonstra-
tive evidence,99 as well as some of the definitional and conceptual
shortcomings of the leading scholarly works on the subject.100 Unfor-
tunately, their article predates Daubert and the subsequent revolution
in expert evidence. However, they do acknowledge the "ever-increas-
ing emphasis in modern trials on lay and expert opinion testimony" and
the resulting fact that "modern lawyers are increasingly turning
to demonstrative evidence to make these opinions understandable to
triers of fact."101 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."102

A major shortcoming of the existing rules of evidence is a failure to
provide an explicit avenue for the admissibility of demonstrative evi-
dence.103 Indeed, strictly speaking, such demonstrative evidence ordi-
narily does not prove or disprove facts of consequence and, therefore,
cannot meet the definition of "relevant evidence" under Rule 401.
Thus, Brain and Broderick advocate modifying Rule 401 to provide

97. Mark Dombroff, Dombroff on Demonstrative Evidence, §§ 1.14-1.15, at 13-
14; § 8.10, at 180-81 (1983). See also Mark A. Dombroff, Innovative Developments
in Demonstrative Evidence Techniques and Associated Problems of Admissibility,
98. 22 Charles Alan Wright & Kenneth W. Graham, Jr., Federal Practice and
Procedure § 5163, at 37 (1st ed. 1978).
100. Id. at 1004-13.
101. Id. at 964.
102. Id. at 968-69.
103. Id. at 1018.
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." Unfortunately, despite the ever-increasing importance of such evidence, it appears that their recommendations have been largely ignored by the Federal Rules Advisory Committee and courts.

D. Experimental Evidence and the Federal Rules

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 applied constraints contained in the existing Federal Rules to curtail the inappropriate use of tests or experiments as demonstrative evidence by experts.

1. Rule 702

The admissibility of testing performed by expert witnesses is governed by Rule 702. That rule requires any expert evidence, whether testimonial or demonstrative, to be 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 Rule 702, the Court in Daubert abandoned the Frye test, a pre-Federal

104. Id. at 1023.
106. See e.g., May Dep't Stores Co. v. Runge, 241 F. 575, 579 (8th Cir. 1917) (quoting Darling v. Westmoreland, 52 N.H. 401, 404 (N.H. 1872)).
107. 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.
108. See text of Rule 702, supra note 3.
Rules standard for the admissibility of novel scientific evidence. The Court replaced *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.\(^{111}\)

*Daubert* explicitly stated that “[the Court’s] discussion [was] limited to the scientific context because that is the nature of the expertise offered here.”\(^{112}\) In the following decade, lower courts reached differing conclusions as to whether those criteria applied solely to scientific testimony, solely to *novel* scientific testimony, or to *all* expert testimony.\(^{113}\) However, a few courts quickly concluded that many of the same hallmarks of reliability should also extend to the testimony of technical experts. A forceful early example was *Stanczyk v. Black & Decker, Inc.*,\(^{114}\) in which the district court excluded the testimony of an engineer who failed to test his theory about the utility of his proposed guard for a miter saw.\(^ {115}\) Judge Zagel illustrated the issue by contrasting the theories of Seventeenth-Century astronomers Tycho Brahe and Galileo—Brahe theorizing without attempting to confirm his theories through observation, and Galileo testing and refuting long-held beliefs through empirical testing.\(^ {116}\) Judge Zagel therefore posed the evidentiary question as: “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* . . . .”\(^ {117}\)

A series of subsequent Seventh Circuit Court of Appeals 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.\(^{118}\) In another case, the court concluded that the admissibility of an engineer’s expert opinions should be based on considerations such as, “[w]hat tests do engineers use to resolve questions of the kind [this expert] addressed? What tests should he have performed? What data did he overlook?”\(^ {119}\) Professor Faigman correctly predicted the ultimate outcome of the issue,

\(^{111}\) *Daubert*, 509 U.S. at 592-94.
\(^{112}\) *Id.* at 590 n. 8. See also *id.* at 600 (Rehnquist, J., concurring) ("Does all this *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’?”).
\(^{115}\) *Id.* at 566-67.
\(^{116}\) *Id.* at 566.
\(^{117}\) *Id.*
\(^{118}\) *Cummins*, 93 F.3d at 362.
\(^{119}\) *DePaepe v. Gen. Motors Corp.*, 141 F.3d 715, 720 (7th Cir. 1998).
stating that "[t]he guiding principle . . . must be Daubert's expectation that hypotheses that can be tested will be tested."120

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 other specialized expert testimony—not just scientific evidence.121 In 2000, Kumho's holding was codified by amending Rule 702 to its current form.122

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.123 Although a few courts 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.124 A flurry of recent cases confirm that, as Professor Feigman asserted over a decade ago, when an expert's hypothesis can be tested, it ordinarily must be tested to be admissible.125

However, just as Rule 702 requires experts to test their hypotheses if possible, it also requires tests to be properly conducted and reliably

---

122. FED. R. EVID. 702. See the text of Rule 702, as amended in 2000, supra note 3. 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]hether 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.").

FED. R. EVID. 702 advisory committee's note.
124. Id.
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. Likewise, Rule 702 has been applied to exclude an expert's tests and the opinions based upon the tests, where the tests were not shown to be the product of a scientifically valid methodology.

2. Rule 403

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." 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 of concern to advocates of the benefits of demonstrative evidence." In one early example, 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.

129. See supra note 8.
132. In re Air Crash Disaster at John F. Kennedy Int'l Airport on June 24, 1975, 635 F.2d 67, 72–73 (2d Cir. 1980).
Likewise, Rule 403 has been applied to exclude experimental evidence when tests performed by experts were found to be misleading.\textsuperscript{133} Indeed, two decades before it decided \textit{Muth}, the same court addressed a similar issue involving the use of a crash test and decided the issue based on Rule 403.\textsuperscript{134} 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.\textsuperscript{135} 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.\textsuperscript{136} 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.\textsuperscript{137} Surprisingly, the court in \textit{Muth} did not cite the earlier Fifth Circuit case.

A number of other courts have applied Rule 403 to determine the admissibility of tests offered to prove the party’s theory of an accident, where the proponent omitted important variables.\textsuperscript{138} For example, in \textit{Chase v. General Motors Corp}.,\textsuperscript{139} the plaintiff alleged that the crash occurred because of a defect in the braking system; GM countered that the crash was unavoidable.\textsuperscript{140} The plaintiff offered tests conducted by NHTSA and other tests by its own expert to show that, absent a defect, the brakes were capable of stopping the car in time to avoid the crash.\textsuperscript{141} However, the tests “were performed in daylight, on flat, straight surfaces under controlled protocols at test facilities or on an

\begin{flushleft}
\par
\textsuperscript{134} Shipp v. Gen. Motors Corp., 750 F.2d 418, 427 (5th Cir. 1985).
\textsuperscript{135} Id.
\textsuperscript{136} Id. See also Gaskell, 985 F.2d at 1062 (“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.”).
\textsuperscript{137} Shipp, 750 F.2d at 427 (“A central tenet of GM’s argument, however, was that all rollover accidents are 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 TransAm involved in the [plaintiff’s accident].”).
\textsuperscript{138} See, e.g., Chase v. Gen. Motors Corp., 856 F.2d 17 (4th Cir. 1988); Gladhill v. Gen. Motors Corp., 743 F.2d 1049 (4th Cir. 1984).
\textsuperscript{139} 856 F.2d 17 (4th Cir. 1988).
\textsuperscript{140} Id. at 18.
\textsuperscript{141} Id. at 18–19.
\end{flushleft}
airport runway.” 142 “The tests used experienced drivers at regulated vehicle speeds and who steered straight ahead.” 143 “Brake pedal force in the NHTSA tests was mechanically applied to produce single axle lock.” 144 “The NHTSA tests were performed on three different surface types under wet conditions.” 145 In the expert’s tests, he applied the brakes himself, “similarly test[ing] the vehicle on both wet and dry surfaces.” 146

The conditions on the night of the crash were significantly different from those in the tests. 147 The plaintiff “was driving downhill on a slight curve to the left and was so turning his car at the time of the accident.” 148 “[T]he surface of the road was covered with snow and ice.” 149 “Instead of applying measured or pre-ascertained brake pressure, [the plaintiff] instinctively applied the brakes.” 150 The Fourth Circuit Court of Appeals analyzed the admissibility issues in a manner that conflated the 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.” 151 It further held that the trial judge should have excluded both the evidence of the tests themselves, as well as the videotape of the tests. 152

In another case, a plaintiff was injured when his vehicle struck a moose on a highway. 153 The district court excluded a video of an inverted drop test of a vehicle similar to the one in which the plaintiff was injured. 154 The plaintiff offered the test to show that the roof of the vehicle lacked sufficient strength. 155 The court found that the test would be misleading to the jury and was therefore inadmissible under Rules 701, 702, and 403. 156 There was no evidence that the magnitude of the forces in the drop test and the moose collision were comparable to each other. 157 The forces imposed during the drop test were vertical, whereas the vehicle was moving forward when it struck the

142. Id. at 19.
143. Id.
144. Id.
146. Id.
147. Id.
148. Id.
149. Id.
150. Id.
152. Id. at 20 (citing Gladhill, 743 F.2d at 1052).
154. Id. at 187–88.
155. Id. at 188.
156. Id. at 190.
157. Id. at 188.
mooose.158 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.159 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 the plaintiff’s collision with the moose.160 Thus, the court was able to resolve the admissibility issue by reference to Rules 702 and 403 rather than 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.161

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. . . .”162 This rule can constrain 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 inadmissi-

159. Id. at 189–90.
160. Id. at 190.
161. Contra Finchum v. Ford Motor Co., 57 F.3d 526, 530 (7th Cir. 1995) (“The sled test was just similar enough to the Finchums’ accident to confuse the jury and leave jurors with the prejudicial suggestion that the Finchums flipped over backwards during the crash.”). Finchum, like Muth, tried to merge the substantial similarity standard with the criteria of Rule 403, without differentiating between the two.
162. 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.

FED. R. EVID. 703.
Rule 703 permits the expert to render the opinion but does not permit the expert to bring 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 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 opinion implies that the experts were permitted to testify about the tests and were merely precluded from showing any test films or photographs to the jury.

---

163. Under earlier law, some courts interpreted Rule 703 to allow the expert to disclose hearsay "for the limited purpose of explaining the basis for his expert opinion." Paddack v. Dave Christensen, Inc., 745 F.2d 1254, 1262 (9th Cir. 1984). Also compare Ronald L. Carlson, *Policing the Bases of Modern Expert Testimony*, 39 Vand. L. Rev. 577, 585–86 (1986) ("To protect against litigation based on unsworn allegations contained in the report of a nontestifying expert, it may be time to consider careful revision of Federal Evidence Rule 703.") with Paul R. Rice, *Inadmissible Evidence as a Basis for Expert Opinion Testimony: A Response to Professor Carlson*, 40 Vand. L. Rev. 583, 584 (1987) ("[T]he introduction of the inadmissible facts or data upon which experts rely no more violates the hearsay rule's spirit than do the volumes of evidence that regularly are introduced through the numerous hearsay rule exceptions.").

164. *Id*. See United States v. Leason, 453 F.3d 631, 638 (4th Cir. 2006) (upholding determination that probative value of hearsay relied upon by expert substantially outweighed prejudicial effect); Turner v. Burlington N. Santa Fe R.R. Co., 338 F.3d 1058, 1062 (9th Cir. 2003) (upholding exclusion of hearsay relied upon by expert because its probative value did not outweigh its prejudicial effect).

165. Fed. R. Evid. 703.

166. *But see* Montgomery v. Mitsubishi Motors Corp., 448 F. Supp. 2d 619, 633 (E.D. Pa. 2006) (testifying expert could offer an opinion based on tests conducted by a non-testifying expert, but 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) (holding that the testifying expert's statements regarding conclusions derived from tests conducted by a non-testifying expert were inadmissible as hearsay).

4. Rule 701

Because so many of the recent cases arise from evidence of tests performed by experts, it is easy to overlook the fact that experimental evidence can be offered by lay witnesses as well. Indeed, many of the older cases arise from testimony of lay witnesses who "tested" some aspect of a party's theory. Therefore, even if Rule 702's reliability requirements obviate the necessity of retaining the substantial similarity standard for experts, one might argue 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 adds 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 Rule 701 should not be used to evade Rule 702. Indeed, even before the 2000 amendments, one circuit had reached a similar conclusion under the earlier version of Rule 701. The commentary and case law surrounding Rule 701 now protect against the use of lay witnesses to obtain admission of otherwise unreliable or prejudicial experiments or tests.

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

168. 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.

FED. R. EVID. 701.

169. See, e.g., May Dep't Stores v. Runge, 241 F. 575 (8th Cir. 1917); see also discussion supra Section II.B.1.

170. FED. R. EVID. 701 advisory committee's note.


reach its own conclusions, the court may exclude as unhelpful the lay opinion of a witness not qualified as an expert.\(^\text{173}\)

E. Substantial Similarity: The Goldilocks Principle?

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 that experts test their testable hypotheses, that the court act as a gatekeeper to ensure the reliability of the test method employed by the expert,\(^\text{174}\) and that the court determine that the opinions are derived reliably from the facts.\(^\text{175}\) Rule 403 enables the court to exclude or otherwise limit relevant evidence if its relevance is substantially outweighed by the danger of unfair prejudice,\(^\text{176}\) confusion of the issues,\(^\text{177}\) or is misleading\(^\text{178}\) or cumulative,\(^\text{179}\) and Rule 703 presumptively excludes backdoor hearsay upon which the expert relied. If any additional legitimate purpose is served by superimposing a substantial similarity requirement on demonstrative evidence that is otherwise reliable and relevant, the cases have failed to identify it. Instead, some courts, such as the Fifth Circuit in \textit{Muth}, persist in applying 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-Federal Rules 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 at least five reasons that continued use of such a standard is not a harmless error.

\begin{enumerate}
\item \textit{Distraction from More Appropriate Standards}

The substantial similarity standard can distract 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,\(^\text{180}\) or a decision to exclude all demonstrative test-

\begin{itemize}
\item \textsuperscript{174} Schwab v. Nissan N. Am., 502 F. Supp. 2d 980, 982 (E.D. Mo. 2007).
\item \textsuperscript{175} \textit{Id.} at 985 (citing Gen. Elec. Co. v. Joiner, 522 U.S. 136, 146 (1997)).
\item \textsuperscript{176} United States v. Gaskell, 985 F.2d 1056 (11th Cir. 1993).
\item \textsuperscript{177} Daskarolis v. Firestone Tire & Rubber Co., 651 F.2d 937, 940 (4th Cir. 1981).
\item \textsuperscript{178} Finchum v. Ford Motor Co., 57 F.3d 526, 530 (7th Cir. 1995).
\item \textsuperscript{179} Mittlieder v. Chicago & Nw. Ry. Co., 441 F.2d 52, 55 (8th Cir. 1971).
\item \textsuperscript{180} Shipp v. Gen. Motors Corp., 750 F.2d 418 (5th Cir. 1985).
\end{itemize}
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; subsequently, the appellate court, 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 quandary deciding whether it supplants applicable provisions of the Federal Rules or, if not, how it fits within those provisions.

By contrast, under Rules 403, 611, and 105, the court can—and should—consider the degree to which such evidence is likely to cause confusion, open up time-wasting collateral issues, produce such a strong visceral response in jurors that it could overwhelm their objectivity, or whether the dissimilarities are explainable. These rules also give the court latitude to 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. 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.182 This was consistent with the Advisory Committee Note to Rule 403, which states that, in deciding “whether to exclude on grounds of prejudice, consideration should be

---

182. *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 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 & Com. 139, 146 (1979) (discussing a cautionary instruction used in admitting a model of an airport in a trial arising from an air crash disaster).
given to the probable effectiveness or lack of effectiveness of a limiting instruction." 183

In Muth, although the Fifth Circuit alluded to the potentially misleading nature of the tests, it decided the case without any explicit analysis of any of the applicable rules of evidence. With one insignificant exception, the Fifth Circuit's entire discussion of the inadmissibility of the Malibu and CRIS tests failed even to cite any Federal Rule. 184 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. 185

2. Vagueness of "Substantial Similarity"

Given the impossibility of conducting a test that is identical to the accident, how can courts make reasoned and logically consistent judgments about how similar a test must 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. 186 How else to account for the fact that two courts reached precisely the opposite conclusion on

183. 22 CHARLES ALAN WRIGHT & KENNETH W. GRAHAM, JR., FEDERAL PRACTICE AND PROCEDURE § 5211 at 244 (1st ed. 1978). See also supra text accompanying note 136.

184. 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 at 566 n.18.


186. 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 added) (citations omitted) with 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 Sharon Panian, Comment, Truth, Lies, and Videotape: Are Current Federal Rules of Evidence 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 the vehicle being tested be "virtually identical" with accident vehicle). But see In re Jacoby Airplane Crash Litigation, 2007 WL 3243920, *3 (D.N.J. Nov. 1, 2007) (distinguishing between a demonstration that "recreates" the accident and one that is offered to "illustrate" the expert's opinion).
the admissibility of the same test in the same kind of case? The same Malibu test video, excluded in Muth, was admitted in Harvey, even though both cases presented rollover product liability claims based on allegations of inadequate roof strength. The imprecision of the term, "substantially similar," is compounded by the broad discretion appellate courts afford the trial court in making its determination.

No doubt, similarities and dissimilarities between a test and the actual conditions can be important in determining the reliability of the test under Rule 702, the relevance of the test under Rule 401, and the degree to which the test may be unfairly prejudicial or misleading under Rule 403. But, even though dissimilarities are a part of the analysis, no one can seriously argue that they should 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; 3) whether the dissimilarities 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; 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 proponent is attempting to oversell the test as proving more than it actually does. It is noteworthy that all of these considerations fall within the parameters of Rules 401, 403, 105, 702, and 703.

188. See United States v. Rackley, 742 F.2d 1266, 1272 (11th Cir. 1984); Wagner v. Int'l Harvester Co., 611 F.2d 224, 232 (8th Cir. 1979). The difficulty is further compounded by the fact that the terminology "substantial similarity" is used in a variety of other contexts. See, e.g., Nachtsheim v. Beech Aircraft Co., 847 F.2d 1261, 1268–69 (7th Cir. 1988); Brooks v. Chrysler Corp., 786 F.2d 1191, 1195 (D.C. Cir. 1986) (admissibility of other incidents in product liability case); Arica Inst. Inc. v. Palmer, 970 F.2d 1067 (2d Cir. 1992) (proof of infringement in copyright case); In re Lesker, 939 F.2d 699 (8th Cir. 1991) (classification of creditor claims in Chapter 13 bankruptcy).
189. See Jodoin v. Toyota Motor Corp., 284 F.3d 272, 280 (1st Cir. 2002) (giving examples from prior cases to illustrate its conclusion that "[w]hen the relevant elements are sufficiently similar," other dissimilarities can be adequately dealt with in cross examination) (emphasis added).
190. See supra text accompanying notes 182 and 183.
191. 22 CHARLES ALAN WRIGHT & KENNETH W. GRAHAM, JR., FEDERAL PRACTICE AND PROCEDURE § 5211 at 244 (1st ed. 1978) (quoting Fed. R. Evid. 403 advisory committee's note: "The availability of other means of proof may also be an appropriate factor.").
3. False Dichotomy between "Re-Creations" and "Abstract Principles"

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. The best judicial solution to the evidentiary dispute should be calibrated to accommodate both the probative value of the evidence and exclude or limit the portion of it that is misleading or unfairly prejudicial. The Federal Rules, particularly Rules 403 and 105, provide avenues for doing this; the substantial similarity standard does not.

Moreover, in some cases, there is a continuum rather than a bright line between the "re-creations" and "demonstrations of abstract principles." In such cases, the test at issue is susceptible to a characterization as fitting into either category, or both. Tests often shed light on abstract and case-specific questions alike, 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 recreation. 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. 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 whatever probative value the evidence might have. That sensible flexibility is 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, which party's version of the facts should the expert be permitted

192. 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.

_Fed. R. Evid._ 105.
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? There is no good reason that testing should be considered by the jury only when all parties agree on all material facts. Furthermore, although it may be tactically advantageous to run a test based on the opposing party's version of events, there is no good reason that admissibility of one party's testing should necessarily be governed by adherence to the opposing party's version of disputed facts.

However, Muth held that the trial court had discretion to deem Ford's CRIS test as "prejudicial" because the test's similarities with Ford's theory of the accident meant that the test sought to recreate the accident rather than demonstrate an abstract scientific principle.\(^\text{193}\) In other words, because the test might be perceived as a recreation in some respect, it was inadmissibly misleading because it was too similar.\(^\text{194}\) Such reasoning leads to a nonsensical analysis because any demonstration of any abstract principle is subject to the same attack as long as the abstract principle has enough overlap with the facts of the disputed accident to be relevant at all. That the CRIS test offered in Muth was consistent with 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. In a case arising from the rollover of a Ford Crown Victoria, can one seriously argue that a rollover test involving a Volkswagen Beatle would be better than one using a similar Crown Victoria?

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 whenever it is substantially similar in any respect.\(^\text{195}\) Thus, if a plaintiff was injured by a 20-pound object falling

---

\(^\text{193}\) Muth v. Ford Motor Co., 461 F.3d 557, 567 (5th Cir. 2006).

\(^\text{194}\) Id.

\(^\text{195}\) It also presupposes that demonstration of an abstract principle is incapable of being inflammatory, misleading, or unfairly prejudicial. For example, in Wilson v. Piper Aircraft Corp., 577 P.2d 1322, 1331 (Or. 1978), the Court stated:

\[\text{[T]he character of the film itself was such that it ought not to have been shown to the jury. Entitled "Restraints for Survival," the film was obviously intended, and is undoubtedly very effective, as an emotional as well as a factual appeal for the increased use of upper torso restraints for aircraft occupants. The facts and conclusions expressed by the narrator were emphasized in various ways, including the liberal use of dramatic music, repeated views of what appears to be a dead body beside the wreckage of a small airplane, and closeups of dummies used in test crashes, their faces smeared with red to simulate blood. Whatever its merits for other purposes, this film was totally out of place in this trial. Any probative value it had which was not merely cumulative of Dr. Snyder's testimony was outweighed by its frank appeal to the emotions of its audience.}\]
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 on the basis that one of 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. Conversely, the substantial similarity standard is sufficiently nebulous to permit the court to exclude such a test, offered to re-create the accident, merely because the plaintiff was injured by falling granite and the expert used marble. Had the court applied the Federal Rules, rather than the court-created dichotomy between "abstract principles" and "re-creations," it could have addressed whether the test was actually unfairly prejudicial to the proponent.

5. Shifting of the Foundational Burden

The substantial similarity standard confuses the question of which party should bear the foundational burden to establish admissibility. Under the Federal Rules, the proponent of expert testimony has the burden of showing that the evidence is relevant and reliable under Rule 702. 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. By contrast, the substantial similarity standard, as articulated by the Tenth Circuit and generally followed by federal courts since then, places the entire burden on the proponent.

Too often, courts prefer to justify decisions based on old case law, even when the statute or rule governing the issue has changed in the interim. This is such a case. Common-law rules of exclusion should not generally trump the Federal Rules, but courts often struggle in such situations. For example, courts have reached differing conclusions as to whether the common-law collateral source rule divests the trial court of discretion to admit such evidence under Rule 403. The Federal Rules themselves explicitly shift the burden to the proponent to show the absence of unfair prejudice in a few exceptional cir-

198. Jackson v. Fletcher, 647 F.2d 1020, 1027 (10th Cir. 1981).
199. Compare, e.g., Sheehy v. S. Pac. Transp. Co., Inc., 631 F.2d 649, 652 (9th Cir. 1980) (indicating that Rule 403 would not allow the trial discretion in this situation due to common-law collateral source principles), with Savoie v. Otto Candies, Inc., 692 F.2d 363, 371 n.8 (5th Cir. 1982) (stating that discretion arises by way of Rule 403).
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 shift the burden of showing unfair prejudice under Rule 403.

The Tenth Circuit's allocation of the burden for experimental evidence did not result from any analysis of any provision in the Federal Rules. The court justified shifting the foundational burden to 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. Rather, *Navajo Freight Lines* 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 a 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

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 bur-

---

200. See, e.g., FED. R. EVID. 703; see also, United States v. Curtin, No. 04-10632, 2007 WL 1500295, at *8 (9th Cir. May 24, 2007); B.K.B. v. Maui Police Dep't, 276 F.3d 1091, 1104 (9th Cir. 2002) (discussing the relationship between Rule 412(b)(2) and Rule 403); Lesko v. Owens, 881 F.2d 44, 61–62 (3d Cir. 1989) (discussing the relationship between Rule 404(b) and Rule 403).

201. *Jackson*, 647 F.2d at 1027.

202. Id. (quoting *Navajo Freight Lines v. Mahaffy*, 174 F.2d 305, 310 (10th Cir. 1949)).

203. Hisler v. State, 42 So. 692, 695 (Fla. 1906).

204. Id. *Navajo Freight Lines*, 174 F.2d at 310, also cited a 1936 California state appellate court 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. Collins v. Graves, 61 F.2d 1198, 1203 (Cal. Ct. App. 1936).
den of showing prejudice. The courts have given no reason why the burden of showing prejudice should be different for testing evidence than for other kinds of evidence.

There might well have been reasons for imposing the entire burden on the proponent a century ago, when the Florida court announced such a rule. At that time, without pretrial discovery and expert disclosure requirements, the proponent could ambush the other party with an experiment. It undoubtedly seemed more sensible to make the proponent shoulder the entire foundational burden than to require the opponent to respond to the evidentiary ambush. Of course, the Federal Rules of Civil Procedure have long since eliminated trial by ambush, and the detailed expert disclosures, expert depositions, motions in limine, and fact discovery available today render any such concern obsolete.

In short, Muth exemplifies the shortcomings of the substantial similarity standard and demonstrates why the First Circuit in Fusco was right in describing the existing state of the law as "muddled." The trial court in Muth rejected the tests Ford offered as "not quite similar enough," yet too similar. 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. Furthermore, the vehicles used in the CRIS test were the same vehicle model (Ford Crown Victoria) as that involved in the case, albeit a newer model year. 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 it justify also excluding the video of the Malibu test, which was conducted using a completely different

205. United States v. Tse, 375 F.3d 148 (1st Cir. 2004).
209. In an analogous vein, in Jones v. Otis Elevator Co., 861 F.2d 655 (11th Cir. 1988), the court questioned the need to continue to observe the common-law rule that a when party fails to produce a witness under its control, the jury may infer that the testimony of the witness would be unfavorable. Because modern federal procedural and evidentiary rules enable discovery of witnesses employed by the opposing party and no longer require a party to "vouch" for the witnesses it calls to testify, the Eleventh Circuit questioned the wisdom of permitting the missing witness charge when the potential witness was within the subpoena power of both parties and physically available to be called. Id. at 659 n.4.
211. Muth v. Ford Motor Co., 461 F.3d 557, 566 (5th Cir. 2006).
212. Id. at 567.
214. Muth, 461 F.3d at 567.
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 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 in advance 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, if the court was applying the substantial similarity standard consistently, it should have excluded all evidence of testing, not just the demonstrative evidence. Indeed, if the court had applied the Federal Rules, 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, because it permitted the jury to hear the opinions of Ford's expert but not to see any of what the court characterized dismissively as "his visual aids."

Undoubtedly, there may be circumstances under which a court should admit evidence of the test but exclude the demonstrative evidence. But such a decision could be made more rationally if based explicitly upon application of Rules 403, 611(a), 703, and 105 rather than the substantial similarity standard.

F. A Modest Proposal: Apply the Rules

As previously suggested, the Federal Rules of Evidence already provide the necessary tools for assessing the admissibility of experimental evidence. They offer a more detailed and systematic methodology for identification and analysis of the pertinent issues relating 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 to make 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 de-

---

215. *Id.* at 566–67.
216. *Id.*
monstrative evidence of testing should be to follow the applicable Federal Rules systematically by considering:

1) Whether the proponent has demonstrated that the tests were reliable and relevant under Rules 702 and 401. In the case of monstrative evidence, "relevant" ought to be understood to mean that they fairly and accurately explain, illustrate, or clarify, other admissible evidence.217

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

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

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

Reconciliation of the applicable provisions of the Federal Rules with a common-law standard that predates the Federal Rules is a well-defined process. Rule 402 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."218 It follows that common-law evidentiary decisions do not trump the Federal Rules.

Daubert presents a case in point. Some courts continued to apply Frye even after adoption of the Federal Rules.219 In Daubert, the Supreme Court was asked to reconcile the pre-Federal Rules Frye test with Rule 702.220 The Court noted that the legislatively-enacted Federal Rules should be interpreted in the same manner as any statute.221 Citing United States v. Abel,222 the Court concluded that "the Rules occupy the field," but that "the common law nevertheless could

217. Brain & Broderick, supra note 89, at 923.
220. Id. at 587.
221. Id. (citing Beech Aircraft Corp. v. Rainey, 488 U.S. 153, 163 (1988)).
serve as an aid to their application."223 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."224 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."225 Because nothing in the Federal Rules incorporated the 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 Frye test was "unconvincing."226 Accordingly, the Court abandoned Frye and, instead, established a new standard consistent with the language and purposes of the Federal Rules.227

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, and the Rules provide no basis for continuing to treat experimental evidence as a discrete category of evidence. The reliability, relevance, and prejudice provisions of the Federal Rules encompass the issues raised by such evidence. It follows that, as with Frye, the pre-Federal Rules substantial similarity standard should give way to the standards already incorporated within the Rules.

III. CONCLUSION

The substantial similarity standard for evidence of tests, like the Frye standard, is an artifact of pre-Federal Rules evidence law. Characterizing such evidence as "experimental," like Frye's treatment of "novel scientific" evidence, should no longer trigger special rules of admissibility. In a post-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. Undoubtedly, the Rules would benefit from clarification of the standards for admissibility of demonstrative evidence, in order to

---

224. Id. at 588.
225. Id.
226. Id. at 589.
227. Another example of a common-law evidence rule supplanted by the Federal Rules of Evidence is the doctrine of res gestae. F.D.I.C. v. Fid. & Deposit Co. of Md., 45 F.3d 969, 979 (5th Cir. 1995). This "obsolete" doctrine was described by other federal courts as "useless, harmful, and almost inescapable of a definition." Stephens v. Miller, 13 F.3d 998, 1003 (7th Cir. 1994) (citing Williams v. Melton, 733 F.2d 1492, 1494 (11th Cir. 1984); Wheeler v. United States, 211 F.2d 19, 23 n.11 (D.C. Cir. 1953)) (en banc).
bring the evidence more clearly under the umbrella of the existing rules, such as was proposed by Brain & Broderick.

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

[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 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.228

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 can distract 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.

228. 2 WIGMORE ON EVIDENCE § 445 (James H. Chadbourn ed. 1979).