

# Defining the "Task at Hand": Non-Science Forensic Science After *Kumho Tire Co. v. Carmichael*

D. Michael Risinger\*

## I. *The Lessons of Kumho Tire Co. v. Carmichael*

The 1970s and early 1980s were a period of virtually unbridled expansion of asserted expertise in civil and criminal courtrooms, limited only by the imagination of an attorney with a point to prove and a hole in her more conventional evidence.<sup>1</sup> The appeal of using such experts stemmed in large part from two aspects of the law, one in regard to experts and one in regard to sufficiency of evidence. Courts allowed experts to phrase opinion testimony in terms of the ultimate issues in the case.<sup>2</sup> If the "opinion" of the expert was competent, the jury might adopt the opinion *in toto*, making failure of proof on the issue legally impossible. Combine this situation with decidedly lax threshold standards of admissibility for expertise, and the stage was set for the acceptance of some fairly questionable practices in the utilization of expertise by litigants. Consequently, although all sides were free to play the game, the result was generally much more favorable to parties with the proof burdens (generally civil plaintiffs and the prosecution in criminal cases, though criminal defendants were substantial players in regard to various affirmative defenses).

In the mid and late 1980s, critics raised their voices in protest, saying that the kind of expertise the courts regularly accepted as admissible was frankly "junk" of scandalous lack of dependability. Voices protested the lack of reliability in both criminal and civil spheres,<sup>3</sup> but the voice that finally spoke

---

\* B.A., Yale, 1966, J.D., Harvard, 1969; Professor of Law, Seton Hall University School of Law. The author would like to thank Michael J. Saks and Mark P. Denbeaux for helpful comments on earlier drafts of this Article.

1. See Richard H. Underwood, "X-Spurt" Witnesses, 19 AM. J. TRIAL ADVOC. 343, 345-48 (1995) (discussing increasing role of expert testimony).

2. See FED. R. EVID. 704(a) (stating that "testimony . . . is not objectionable because it embraces an ultimate issue to be decided by the trier of fact").

3. See D. Michael Risinger et al., *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification "Expertise,"* 137 U. PA. L. REV. 731,

loudest and was heard most clearly, spoke almost exclusively of the injustice of junk expertise used against civil defendants. I refer, of course, to Peter Huber and his 1991 book, *Galileo's Revenge*,<sup>4</sup> which popularized the phrase "junk science." Given the polemical success of that book, it seems unlikely to have been pure coincidence that the United States Supreme Court chose a civil case to review the appropriate threshold criteria of reliability for expert testimony, or that its two subsequent forays into these waters have also been in civil cases. Be that as it may, the pronouncements of the Supreme Court are given as trans-substantive constructions of the Federal Rules of Evidence, and so have application in criminal as well as civil cases.

The wellspring case, as everyone knows, is *Daubert v. Merrell Dow Pharmaceuticals, Inc.*<sup>5</sup> Like many groundbreaking decisions, *Daubert* was neither fully worked out nor fully coherent. The day after the Court decided *Daubert*, the *Washington Post* characterized it as a victory for those who wanted expertise more easily admitted, while the *New York Times* characterized it as a victory for those who wanted more expertise rejected.<sup>6</sup> This schizoid characterization of the case has continued in both academic commentary and lines of judicial decision down to the present time.<sup>7</sup> In a recent opinion, Judge Gertner of the Massachusetts U.S. District Court attributes this to what she calls "competing vectors" in the *Daubert* opinion.<sup>8</sup> The first

---

779 (1989) (claiming that "the law does not yet know how to deal with science, or with things asserting themselves to be science").

4. PETER W. HUBER, *GALILEO'S REVENGE* (1991).

5. 509 U.S. 579 (1993).

6. Compare Joan Biscupic, *Judges Get Broader Discretion in Allowing Scientific Testimony*, WASH. POST, June 29, 1993, at A6 (stating that Supreme Court relaxed standards for admission of scientific testimony in *Daubert*), with Linda Greenhouse, *Justices Put Judges in Charge of Deciding Reliability of Scientific Testimony*, N.Y. TIMES, June 29, 1993, at A13 ("The 7-2 decision invited judges to be aggressive in screening out ill founded or speculative scientific theories.").

7. Compare David L. Faigman et al., *Check Your Crystal Ball at the Courthouse Door, Please: Exploring the Past, Understanding the Present, and Worrying About the Future of Scientific Evidence*, 15 CARDOZO L. REV. 1799, 1801-02 (1994) (arguing that *Daubert* generally raises bar on admissibility) with Arvin Maskin, *The Impact of Daubert on the Admissibility of Scientific Evidence: The Supreme Court Catches Up With a Decade of Jurisprudence*, 15 CARDOZO L. REV. 1929, 1942 (1994) (claiming that *Daubert* should be viewed as making admission easier). See also 1 DAVID L. FAIGMAN ET AL., *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* § 1-3.3, at 17-28 (2000 Supp.) (comparing *Daubert* test to *Frye* test). Faigman et al. make the most persuasive analysis. They argue that *Daubert* is more limiting than the *Frye* general acceptance test when the *Frye* test would let in expertise generally accepted by a "community" on shaky grounds, but less limiting than *Frye* when something too new to be said to have gained "general acceptance" can nevertheless be shown to be extremely reliable. *Id.*

8. See *United States v. Hines*, 55 F. Supp. 2d 62, 65-67 (D. Mass. 1999) (analyzing competing interpretations of *Daubert*).

vector, which points towards a more rigorous standard of reliability, is characterized by the Court's emphasis on scientific standards and its encouragement of gatekeeping review under Rule 702 of the Federal Rules of Evidence.<sup>9</sup> The second vector, which points in the opposite direction, is characterized by the uniqueness of the trial setting, the "assist the trier" standard,<sup>10</sup> and flexibility (coupled with its rejection of "general acceptance" as an absolute *sine qua non* of admissibility for scientific expertise).<sup>11</sup>

Even as it became reasonably clear that the effect of the *Daubert* decision in regard to scientific testimony was to raise the bar for admission,<sup>12</sup> two general schools of thought about the "true meaning of *Daubert*" in regard to "non-scientific" expertise emerged. The first school saw *Daubert* as essentially a general construction of Rule 702 and the judge's systemic gatekeeping duties in regard to the sufficient reliability of all proffered expert testimony. To members of this school, *Daubert*'s particular expositions about scientific evidence were important as guides to the kind of reliability that ought to be required of all expertise, even if the so-called "*Daubert* factors" (which the *Daubert* opinion itself said were neither *sine qua nons* themselves, nor exhaustive) applied most powerfully to the products of the conventional sciences. People of this persuasion have, under the banner of *Daubert*, tended to call upon courts to examine proffered claims of expertise specifically and critically and have tended to advocate for generally rigorous standards of reliability as a condition of admissibility.

The other main school of thought believed that *Daubert* ought to be read as limited to scientific expertise, narrowly confined to the experimental sciences. As to all other forms of expertise, especially expertise with a claimed "experiential" or "clinical" component, this school of thought understood *Daubert*'s broader references to Rule 702 as no more than restating the pre-existing understanding of the duty of the court under the "helpfulness" standard, without suggesting that this standard ought to be tightened up regarding reliability. Thus, as to non-scientific evidence, the second school of thought regarded practice under Rule 702 as unaffected and unchanged. To the extent some explicit approach to reliability was thought necessary for such testimony, people of this persuasion have tended to favor a more general or

---

9. *Id.*

10. This standard is named after its emphasis on the language in Rule 702 that expert testimony need only "assist the trier of fact" to be admissible.

11. *Hines*, 55 F. Supp. 2d. at 66-67.

12. In a recent article, I examine more than 2000 cases that have cited *Daubert* from its decision until August 2000. This study reveals that most reported *Daubert* challenges have been in civil cases, most have been made against civil plaintiffs by defendants, and most have been successful. See generally D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALB. L. REV. 99 (2000).

global examination of the claimed abilities of practitioners of the asserted expertise. They have advocated, at least where applicable, some version of either a "sufficient experience" test<sup>13</sup> (relying on the expert's previous more or less similar experience, without further proof that the experience has resulted in any reliable skill), or a "guild" test,<sup>14</sup> in which the existence of an organized

---

13. See Edward J. Imwinkelried, *The Next Step After Daubert: Developing a Similarly Epistemological Approach to Ensuring the Reliability of Non-Scientific Testimony*, 15 CARDOZO L. REV. 2271, 2292-94 (1994) (laying out "sufficient experience" approach); see also Lisa M. Agrimonte, *The Limitations of Daubert and Its Application to Quasi-Scientific Experts, A Two Year Case Review of Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 113 S. Ct. 2787 (1993), 35 WASHBURN L.J. 134, 152-56 (1995) (setting forth new two-factor test for admissibility of quasi-scientific expert testimony); L. Timothy Perrin, *Expert Witness Testimony: Back to the Future*, 29 U. RICH. L. REV. 1389, 1457-62 (1995) (arguing for new test for reliability requirement).

To be fair to Professor Imwinkelried, the test was set out as a preliminary step in an article pointing out the difficulties of formulating reliability tests for non-scientific expertise, especially "clinical" or "experience-based" expertise. Professor Imwinkelried himself well understood that *Daubert's* general aspects required gatekeeping vigilance as to all expertise. Unfortunately, the approach he set out was easily embraced by proponents of the pre-*Daubert* status quo because all that it required was the testimony of the witness that he had had lots of experience, and that much of it was in circumstances substantially like those in the case at hand. See generally, e.g., testimony of Grant R. Sperry, *infra* note 107. Practitioners of all sorts of questionable claimed skills can pass this test. It is not that the test is not sufficient for some kinds of expertise in some circumstances. It works fine for what I have called "everyday summarizational" experts, who testify to such things as industry practice from their years in an industry. See D. Michael Risinger, *Preliminary Thoughts on a Functional Taxonomy of Expertise*, in 3 MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY § 34-2.1, at 4-7 (David L. Faigman et al. eds., 1999) (discussing summarizational experts). The problem is that it supplies little in the way of validation for what I have called "translational" experts, such as handwriting identification experts, who claim that they can translate their experience into particular adjudicative inferences. *Id.* at 8-14. Here, we must worry about the reliability of not only the subjective data base, but also the subjective translational system applied to it. In short, merely showing up at the scene after auto accidents, even hundreds of times, is a weak warrant to believe a witness's inferences about what happened in the accident itself.

14. The "guild" test goes beyond the "sufficient experience" test by focusing inquiry on the existence of a group that certifies training, experience, and methodology. Acceptance by such a group establishes reliability. Note that this is not the *Frye* test applied to non-scientific expertise because in the absence of such group acceptance an individual witness might be found reliable for other reasons. Like the *Frye* test, however, acceptance by such a group guarantees admissibility. The problem, of course, is that astrology can pass this test. For the most extended and explicit assertion of the guild approach, see Andre A. Moenssens, *Handwriting Identification Evidence in the Post-Daubert World*, 66 UMKC L. REV. 251, 291-93 (1998) (arguing that guild approach is preferable test). See also Agrimonte, *supra* note 13, at 155 (arguing for new two-factor test); Daniel J. Capra, *The Daubert Puzzle*, 32 GA. L. REV. 699, 741-46 (1998) (analyzing cases applying guild test); J. Brook Latham, *The "Same Intellectual Rigor" Test Provides an Effective Method for Determining the Reliability of All Expert Testimony, Without Regard to Whether the Testimony Comprises "Scientific Knowledge" or "Technical or Other Specialized Knowledge,"* 28 U. MEM. L. REV. 1053, 1063-68 (1998)

group which supervises accreditation (and an expert's membership in it) is taken as a sufficient warrant to infer reliability for admissibility purposes. In the courts, these usually conflated approaches<sup>15</sup> have been especially prevalent in regard to the products of "forensic science" in criminal cases. Unfortunately for their adherents, the Supreme Court's decision in *Kumho Tire Co. v. Carmichael*<sup>16</sup> has pretty much destroyed the tenability of these approaches.<sup>17</sup>

I could not write a more elegant or efficient summary of circumstances involved in the *Kumho Tire* case than that of the Reporter of Decisions in the Syllabus:

When a tire on the vehicle driven by Patrick Carmichael blew out and the vehicle overturned, one passenger died and the others were injured. The survivors and the decedent's representative, respondents here, brought this diversity suit against the tire's maker and its distributor (collectively Kumho Tire), claiming that the tire that failed was defective. They rested their case in significant part upon the depositions of a tire failure analyst, Dennis Carlson, Jr., who intended to testify that, in his expert opinion, a defect in the tire's manufacture or design caused the blow out. That opinion was based upon a visual and tactile inspection of the tire and upon the theory that in the absence of at least two of four specific, physical symptoms indicating tire abuse, the tire failure of the sort that occurred here was caused by a defect. Kumho Tire moved to exclude Carlson's testimony on the ground that his methodology failed to satisfy Federal Rule of Evidence 702, which says: "If scientific, technical, or other specialized knowledge will assist the trier of fact . . . , a witness qualified as an expert . . . may testify thereto in the form of an opinion." Granting the motion (and entering summary judgment for the defendants), the District Court acknowledged that it should act as a reliability "gatekeeper" under *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, in which this Court held that Rule 702 imposes a special obligation upon a trial judge to ensure that scientific testimony is not only relevant, but reliable. The court noted that *Daubert* discussed four factors – testing, peer review, error rates, and "acceptability" in the relevant scientific community – which might prove helpful in

---

(arguing for application of same test to scientific or technical expert testimony); Peter B. Oh, *Assessing Admissibility of Nonscientific Expert Evidence Under Federal Rule of Evidence 702*, 64 DEF. COUNS. J. 556, 565-67 (1997) (explicitly advocating *Frye* test for non-scientific evidence); Thomas M. Reavley & Daniel A. Petalas, *A Plea for Return to Evidence Rule 702*, 77 TEX. L. REV. 493, 511-13 (1998) (discussing need for clear objective test). The practical (though often inexplicit) adoption of the guild test by courts is discussed *infra* at Section II.

15. In the opinions, one often has to infer the test implicitly being used, usually from the Court's recitation of training, experience, and guild membership, followed by a conclusory declaration of reliability. See *infra* Section II (applying lessons of *Kumho Tire*). When the two tests are hopelessly conflated, I often refer to the mixed product simply as the guild test.

16. 526 U.S. 137 (1999).

17. *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999).

determining the reliability of a particular scientific theory or technique . . . and found that those factors argued against the reliability of Carlson's methodology. On the plaintiffs' motion for reconsideration, the court agreed that *Daubert* should be applied flexibly, that its four factors were simply illustrative, and that other factors could argue in favor of admissibility. However, the court affirmed its earlier order because it found insufficient indications of the reliability of Carlson's methodology. In reversing, the Eleventh Circuit held that the District Court had erred as a matter of law in applying *Daubert*. Believing that *Daubert* was limited to the scientific context, the court held that the *Daubert* factors did not apply to Carlson's testimony, which it characterized as skill – or experience – based.<sup>18</sup>

Thus, *Kumho Tire* squarely presented the issues of the reach of the general *Daubert* approach to threshold reliability under Rule 702 and of *Daubert's* application both to "non-science" and to claims of expertise that were at least in part "clinical."<sup>19</sup> In the first paragraph of the opinion, Justice Breyer discusses *Daubert's* general holding that:

the Federal Rules of Evidence "assign to the trial judge the task of ensuring that an expert's testimony both rests on a reliable foundation and is relevant to the task at hand," and its [discussion of] certain more specific factors, such as testing, peer review, error rates and 'acceptability' in the relevant scientific community, some or all of which might prove helpful in determining the reliability of a particular scientific "theory or technique."<sup>20</sup>

Two sentences later, he declares "[w]e conclude that *Daubert's* general holding – setting forth the trial judge's general gatekeeping obligation – applies not only to testimony based on 'scientific' knowledge, but also to testimony based on 'technical' and 'other specialized' knowledge."<sup>21</sup> Expanding on this point in Part II of the opinion, the Court writes that:

In *Daubert*, this court held that Federal Rule of Evidence 702 imposes a special obligation upon a trial judge to "ensure that any and all scientific testimony . . . is not only relevant but reliable." The initial question before us is whether this basic gatekeeping obligation applies only to "scientific" evidence or to all expert testimony. We, like the parties, believe that it applies to all expert testimony.<sup>22</sup>

---

18. *Id.* at 137-38 (citations omitted).

19. *See id.* at 154 (stating that expert in *Kumho Tire* relied on "visual and tactile inspection" to determine "minute relative shoulder/center treadwear differences," and then determined "that the tire before him had not been abused").

20. *Id.* at 141 (quoting *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 597 (1993)).

21. *Id.*

22. *Id.* at 147 (citations omitted).

Later in the opinion, Justice Breyer states that "*Daubert's* general principles apply to the expert matters described in Rule 702. The Rule, in respect to all such matters, 'establishes a standard of evidentiary reliability.'<sup>23</sup>

Thus, post-*Daubert* decisions that rested on a fundamental distinction between scientific evidence and other kinds of expertise and explicitly or by implication applied a less rigorous standard of reliability to "non-science" have had their main rationale removed and their results at least called into question and put back into play. Yet when these issues are put back into play, how is the game supposed to be played? A court must determine reliability, but does every kind of expert evidence in every context have to meet the same threshold level of reliability to gain admission? The Court does not address this question. Elsewhere, I have argued that there ought to be varying levels of foundational reliability, with that required for prosecution-proffered expertise in criminal cases being very high, especially when it goes to issues of "brute fact" guilt or innocence, such as the identity of the defendant as the perpetrator.<sup>24</sup>

Be that as it may, what is clearly not consistent with *Kumho Tire* is any attempt to approach an issue of reliability globally. That is, reliability cannot be judged globally, "as drafted," but only specifically, "as applied." The emphasis on the judgment of reliability as it applies to the individual case, to the "task at hand,"<sup>25</sup> runs through the opinion like a river. Although the Court does say early in the opinion that "the law grants a district court . . . broad latitude when it decides how to determine reliability"<sup>26</sup> and later says it grants "considerable leeway in deciding in a particular case how to go about determining whether particular expert testimony is reliable,"<sup>27</sup> the Court professes that "a trial court should consider the specific factors identified in

---

23. *Id.* at 149 (quoting *Daubert*, 509 U.S. at 590).

24. Risinger, *supra* note 13, § 34-2.4, at 16.

25. *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 141 (1999). I wish to emphasize that the requirement of task-at-hand analysis is general and applies to all cases, even though I have illustrated it later in the Article with regard to non-science forensic science. In the many, many articles dealing with *Kumho Tire* as of this writing, the centrality of particularized task-at-hand analysis to a proper understanding of the requirements of *Kumho Tire*, and of the limits of judicial discretion under it, generally has not been recognized. In the few articles in which there has been some recognition, it usually has been brief, buried, and unemphasized. See, for example, Michael H. Graham, *The Expert Witness Predicament: Determining "Reliable" Under the Gatekeeping Test of Daubert, Kumho, and Proposed Amendment Rule 702 of the Federal Rules of Evidence*, 54 U. MIAMI L. REV. 317, 339, 341 (2000). The most extensive treatment appears to be in Maj. Victor Hansen, *Rule of Evidence 702: The Supreme Court Provides a Framework for Reliability Determinations*, MIL. L. REV., Dec. 1999, at 1, 41-42, where two paragraphs are devoted to the notion and its implications.

26. *Id.* at 142.

27. *Id.* at 152.

*Daubert* where they are reasonable measures of the reliability of expert testimony."<sup>28</sup> What the Court clearly refers to is the power of the trial judge to select the most appropriate tests of reliability for the application of expertise at issue in the case, not a discretion to treat expertise globally rather than specifically: "[W]hether *Daubert*'s specific factors are, or are not, reasonable measures of reliability in a particular case is a matter that the law grants the trial judge broad latitude to determine."<sup>29</sup>

In fact, in Part I of the opinion, which closely follows the first-quoted passage above, the Court lays the foundation for the particularized approach. The Court's analysis of what the expert was claiming to be able to conclude in the case at hand and the reasons asserted for his claims, is a powerful instance of teaching by example. And it seems clear that the Court intends this lesson because the Court constructs its very particular analysis, not from quotations drawn from the opinion of the District Court, but directly from the record of the expert's deposition and other record documents.<sup>30</sup> In Part I, the Court begins the demonstration of what it wants from district courts in terms of specificity and sophistication of analysis.

If Part I shows how to frame the particular claims of expertise at issue in a particular case, Part III of the opinion illustrates how to apply various factors to an evaluation of the reliability of those claims in the particular case. Indeed, Part III is quite specific in its purpose. "We further explain the way in which a trial judge 'may' consider *Daubert*'s factors by applying these considerations to the case at hand."<sup>31</sup> It is also quite specific in its rejection of a global approach:

For one thing, and contrary to respondents' suggestion, the specific issue before the court was not the reasonableness in general of a tire expert's use of a visual and tactile inspection to determine whether overdeflection had caused the tire's tread to separate from its steel-belted carcass. Rather, it was the reasonableness of using such an approach, along with Carlson's particular method of analyzing the data thereby obtained, to draw a conclusion regarding *the particular matter to which the expert testimony was directly relevant* . . . . The relevant issue was whether the expert could reliably determine the cause of *this* tires separation.<sup>32</sup>

---

28. *Id.*

29. *Id.* at 153. As Justice Scalia points out in concurrence, failure to consider the most relevant criteria of reliability in a given case may be an abuse of discretion. *See infra* note 40 and accompanying text (quoting Justice Scalia).

30. This is not lost on Justice Stevens, who dissents on the basis that the Court should not have undertaken the particularized "well reasoned factual analysis" of Part III of the opinion. Rather, Justice Stevens believes that the Court should have remanded to the Court of Appeals for that purpose. *Kumho Tire*, 526 U.S. at 159 (Stevens, J., dissenting).

31. *Id.* at 153.

32. *Id.* at 153-54 (emphasis added).

And later:

Respondents now argue to us, as they did to the District Court, that a method of tire failure analysis that employs a visual/tactile inspection is a reliable method, and they point both to its use by other experts and to Carlson's long experience working for Michelinas sufficient indication that that is so. But no one denies that an expert might draw a conclusion from a set of observations based on extensive and specialized experience. Nor does anyone deny that, as a general matter, tire abuse may often be identified by qualified experts through visual or tactile inspection of the tire. As we said before . . . the question before the trial court was specific, not general. The trial court had to decide whether this particular expert had sufficient specialized knowledge to assist the jurors "in deciding the particular issues in the case."

The particular issue in this case concerned the use of Carlson's two-factor test and his related use of visual/tactile inspection to draw conclusions on the basis of what seemed small observational differences.<sup>33</sup>

It seems beyond dispute that the "global" approach to establishing reliability is unavailable in any case where a serious challenge to the reliability of a particular application of proffered expertise is raised, that is, any case where its "factual basis, data, principles, methods, or their application are called sufficiently into question."<sup>34</sup> So, it would seem that any post-*Daubert* case that relied on such a global approach would once again have had the ground upon which it stood pulled from beneath it, and its conclusions would, on that basis, be open to question.

Finally, we come to the tenability of the "sufficient experience test" and the "guild test" under the principles and approach of *Kumho Tire*. By definition, any general and exclusive adoption of such tests for all non-scientific evidence becomes untenable once the *Daubert* approach is applied in principle, because both *Daubert* and *Kumho Tire* emphasized the flexible and multi-factorial nature of any defensible evaluation of reliability. Yet both factors may have a legitimate role to play in some kinds of expert situations. Substantial experience of relevant similarity to what is at issue in the case at hand is a necessary condition for the reliability of experience-based expertise, but in most contexts it is not a sufficient condition to establish reliability.

Only when the expert is playing an everyday "summarizational" role is mere experience, no matter how extensive or similar, arguably sufficient.<sup>35</sup> When a witness is called to play a role beyond this and to make conclusions

---

33. *Id.* at 156-57 (citations omitted).

34. *Id.* at 149.

35. A "summarizational" expert is one who is called solely to summarize the net results of previous experience for the jury's education, such as a person with thirty years experience in the shoe business called to testify about industry practices. *See* Risinger, *supra* note 13, § 34-2.1, at 4-7 (defining summarizational experts).

or inferences about adjudicative facts in the case at hand, the testimony is based in part on experience, but in part on some translation scheme to mediate between previous experiences and a particular conclusion in this case. In those circumstances, reliability is dependent on both sufficient experience and a reliable translation system.<sup>36</sup> Perhaps where there are real-world, practice-based, empirically unambiguous indices of success or failure in coming to one's conclusions, we might rationally rely upon experience not only to provide the expert's data base, but also to authenticate the reliability of the conclusory skills involved. This is perhaps the case of Judge McKenna's famous "harbor pilot" example in *United States v. Starzecpyzel*.<sup>37</sup> However, in the case of much of what is called "forensic science," no such unambiguous accuracy feedback exists in normal practice. Unlike the harbor pilot, who either arrives at the right dock or does not, and knows it, a person making a forensic bite mark identification, for example, usually only knows if his conclusion was right or wrong by whether or not a jury agrees with him. In such cases, reliance on the fact that the witness had experience alone as an index of reliability would be irrational. Application of the "astrology test," the litmus test for unacceptable tests of reliability, easily demonstrates this.

The court in *Kumho Tire* explicitly recognizes that astrology is a discipline that "lacks reliability"<sup>38</sup> and that application of a test in a way that resulted in the admission of testimony by such a discipline would be error, using both astrology and necromancy as examples.<sup>39</sup> However, astrologers might pass a simple experience test. They may have dealt with thousands of horoscopes and thousands of cases involving the astrological implications of a particular star chart to honesty. It is not their experience that we doubt, but their methods of conclusion. So, in circumstances when experience alone does not resolve the main doubts about reliability, it would be irrational, and therefore an abuse of discretion to rely upon it. This sort of result was clearly on Justice Scalia's mind when he states in concurrence (joined by Justices O'Connor and Thomas) that he takes the Court's opinion to be clear that:

[T]he discretion [the Court] endorses – trial court discretion in choosing the manner of testing expert reliability – is not discretion to abandon the

---

36. See *id.* § 34-2.3, at 8-10 (discussing witnesses whose role goes beyond mere summarization).

37. See *United States v. Starzecpyzel*, 880 F. Supp. 1027, 1029 (1995) (analogizing handwriting expert to harbor pilot who learns by experience). The problems with the harbor pilot analogy are discussed in D. Michael Risinger & Michael J. Saks, *Science and Nonscience in the Courts: Daubert Meets Handwriting Identification Expertise*, 82 IOWA L. REV. 21, 33-34 (1996), and in Risinger, *supra* note 13, § 34-2.4, at 14.

38. *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 151 (1999).

39. *Id.*

gatekeeping function. I think it is worth adding that it is not discretion to perform the function inadequately. Rather, it is discretion to choose among reasonable means of excluding expertise that is *fausse* and science that is junky. Though, as the Court makes clear today, the *Daubert* factors are not holy writ, in a particular case the failure to apply one or another of them may be unreasonable, and hence an abuse of discretion.<sup>40</sup>

If the sufficient experience test is inadequate in most applications, the guild test is worse because there is no form of expertise for which it is, by itself, an adequate proxy for reliability analysis. Even though it may be that the failure of a particular expert to adhere to the minimum practice standards of a group with which he associates himself might rationally bear on the unreliability of his testimony,<sup>41</sup> adherence to such standards cannot establish reliability when, as is often the case, it is the very reliability of the standard practice that is in issue. The guild test does at least claim to deal with reliability of the process beyond individual experience, but the reliability judgment is delegated to a group that, by definition, already believes in the process. The guild test trades the *ipse dixit* of the individual for the *ipse dixit* of the group.<sup>42</sup> As such, it flunks the astrology test even more dramatically than the experience test because judgments concerning what constitutes sufficiency of experience also are largely abandoned to the guild. There may already be astrology groups in place that could pass any version of the guild approach that did not look at independent evidence of reliability, and there

---

40. *Id.* at 158-59 (Scalia, J., concurring).

41. I take this to be the point of various references to such circumstances in the *Kumho Tire* opinion. Carlson, the expert in that case, claimed to be a member of a subset of the community of engineers which dealt with the examination of failed tires to determine causes of failure, and as the court was at pains to point out, nobody in the case denied that "tire abuse may often be identified by qualified experts through visual and tactile inspection of the tire." *Id.* at 156. The specific issue concerned whether this was one of those times, and the fact that Carlson appeared to be unique, or at least rare, in claiming that the *non-existence* of tire abuse could be established by the methodology he employed was clearly relevant to assessing the reliability of that methodology. Hence, the references to whether an expert's "preparation is of a kind that others in the field would recognize as acceptable," *id.* at 151, that the objective in this case was "to make certain that an expert, whether basing testimony upon professional studies or personal experience, employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field," *id.* at 152, that no other "experts in the industry use Carlson's two factor test," *id.* at 157, and no one "[refers] to any articles or papers that validate Carlson's approach." *Id.* These references show that in the particular circumstances of a case like *Kumho Tire*, deviations from practice standards of a group with which the witness identifies are relevant. They do not establish the acceptability of a guild test when the reliability of the group practice itself is what is being challenged.

42. *See id.* at 157 ("Nothing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert." (quoting *General Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997))).

certainly are such groups of graphologists.<sup>43</sup> As already noted, the majority in *Kumho Tire* recognizes the inadequacy of such "general acceptance" by a community when it is the reliability of the discipline itself, at least in its particular application to the case at hand, that is the issue.

## II. *Applying the Lessons of Kumho Tire Co. v. Carmichael to pre-Kumho Tire Cases*

With these points in mind, we turn to the issues of handwriting identification reliability, as dealt with in the post-*Daubert*/pre-*Kumho Tire* cases and in subsequent cases. Hopefully the reader will recognize that points made in regard to the handwriting have general application to more than merely handwriting. Other topics with applications of controversial reliability include bitemark identification,<sup>44</sup> toolmark identification,<sup>45</sup> arson investigation,<sup>46</sup> and even some aspects of fingerprint identification, firearms identification and forensic pathology. However, I have chosen handwriting identification to illustrate the general problems of the approach of courts to prosecution proffered non-science forensic science for two reasons: first, because it is the area I know best and, more importantly, because it is the area of non-science forensic science that has had the greatest amount of action in the courts in the post-*Daubert* period.

Not that this has been a lot. Since the date of the *Daubert* decision, more than 300 reported cases have noted the presence of questioned document examiner evidence. Nearly 90% of these cases have involved handwriting identification.<sup>47</sup> Yet so far as can be determined from the

---

43. "Graphology" involves the claimed ability to infer personality characteristics from handwriting. It is generally inadmissible in court, but is used by some employers for employee screening. See generally Julie A. Spoh, Note, *The Legal Implications of Graphology*, 75 WASH. U. L.Q. 1307 (1997) (discussing uses of graphology).

44. See 2 FAIGMAN ET AL., *supra* note 7, at § 24 (surveying field of bitemark identification).

45. See *id.* at § 23 (surveying field of firearms and toolmark identification).

46. See *id.* at § 26 (surveying field of arson investigation and identification).

47. The numbers are rounded because it is somewhat difficult to be absolutely sure of exact completeness, and because it is the general magnitude that counts for my purposes anyway. However, anyone interested can examine the cases themselves by doing the following: Go to Westlaw, and on to the Allfeds, Allstates, and Military Cases databases, search for "document examiner," "handwriting identification," and "handwriting analysis" separately since the date of *Daubert*. Go through the cases and subtract those dealing with other aspects of document examination, such as typewriter identification. The number given in the text does not include the substantial number of cases where criminal defendants complain that the prosecution failed to produce document examiner testimony, or where they attacked their own lawyer for ineffective assistance of counsel for so failing.

opinions, litigants have effectively raised dependability issues in only nine of those cases, all in federal court.<sup>48</sup> In the state courts, where the majority of cases arise (split about evenly between civil and criminal cases), not a single reported opinion, with one possible exception,<sup>49</sup> reflects a significant reliability challenge to any aspect of standard document examiner practice.<sup>50</sup> Although in some cases the non-expert evidence is so strong that mounting an attack on the reliability of the expertise might arguably be thought a useless waste of resources, there are other cases, including criminal cases, in which the handwriting identification is central yet no challenge has been mounted.<sup>51</sup> (It does make one wonder what in the sociology of the criminal defense bar accounts for such a result.)

---

48. See *infra* note 52 and accompanying text (listing reported cases raising dependability issues). Two other unreported cases, *United States v. McVeigh*, No. 96-CR-68, 1997 WL 47724 (D. Colo. Trans. Feb. 5, 1997) and *United States v. Brown*, No. CR 99-184 ABC (C.D. Cal. Dec. 1, 1999), have also involved reliability challenges. Judge Matsch's decision in *McVeigh*, which allowed the document examiner to point out similarities between the questioned document and the known exemplars, but not to give a conclusion about the authorship of the questioned document, has been influential even though it was promulgated without formal opinion. It was also influential in the *McVeigh* case itself, since the prosecution chose not to call its document examiner after the court made a ruling limiting his testimony.

*McVeigh* also illustrates the problem of defining what constitutes a "reported opinion" in the days of databases. The oral argument that led to Judge Matsch's decision is reported at 1997 WL 47724. The colloquy, though extensive, does not reveal sufficient facts to determine exactly what task was at issue in the case, beyond the fact that some documents were going to be attributed to McVeigh by a document examiner after comparing them with known samples of McVeigh's handwriting. There may have been an issue of printing comparison, or printing to cursive comparison, but that is not clear. Clearly, Judge Matsch does not formulate the specific task at hand with the particularity required by *Kumho Tire*, but his result has been very influential, as evidenced by its impact on *United States v. Hines* and subsequent cases. See *infra* notes 125-37 and accompanying text (discussing *Hines*); *infra* notes 139-146 and accompanying text (discussing *Santillan*).

49. See *Basinger v. Commonwealth*, No. 2968-98-4, 2000 WL 724037 (Va. Ct. App. June 6, 2000). *Basinger* was a forgery case in which the defense attorney generally objected to the reliability of handwriting identification at trial. The objection was overruled by the trial judge on precedential grounds and disposed of in an equally cursory fashion on appeal.

50. *State v. Wilson*, No. 69346, 1997 WL 127186 (Ohio Ct. App. Mar. 20, 1997), might have involved a general challenge to reliability, but the opinion is too cursory and general to tell. The trial court in *Eubanks v. Hale*, 752 So. 2d 1113 (Ala. 1999), excluded the testimony of a document examiner in a vote challenge case (invoking *Daubert*) because the examiner admitted he had not followed standard procedures in examining the challenged ballot signatures, but even this exclusion was found to be an abuse of discretion on appeal.

51. See, e.g., *State v. Stokes*, 853 S.W.2d 227, 239 (Tex. App. 1993) (upholding conviction partially based on unchallenged testimony of forensic document examiner); *State v. Wilson*, 682 N.E.2d 5, 10 (Ohio Ct. App. 1996) (upholding conviction based in part on handwriting evidence).

Turning now to the reliability challenges reflected in the reported cases, as noted, there are nine such cases, all federal, generating ten opinions.<sup>52</sup> *United States v. Starzecpyzel*<sup>53</sup> is the first handwriting expertise reliability case of the modern era.<sup>54</sup> In that case, Roberta and Eileen Starzecpyzel were charged with having stolen various works of art from Roberta's elderly (and now senile) aunt.<sup>55</sup> They claimed that the paintings were a gift made prior to the aunt's impairment.<sup>56</sup> Part of the evidence against them was the proposed testimony of a questioned document examiner who, after examining numerous authentic signatures of the aunt on checks and other documents, concluded that the aunt's signatures on deeds of gift for the artwork were forgeries.<sup>57</sup> He did not claim to be able to identify either defendant as the forger.<sup>58</sup>

Judge McKenna of the U.S. District Court for the Southern District of New York held an extensive hearing on the state of knowledge concerning the reliability of such asserted expertise. Judge McKenna's opinion examines the claims of handwriting identification expertise to scientific status at length, and rejects them.<sup>59</sup> Having done this, however, he concludes that because such

52. These ten opinions are: *United States v. Paul*, 175 F.3d 906 (11th Cir. 1999); *United States v. Jones*, 107 F.3d 1147 (6th Cir. 1997); *United States v. Velasquez*, 64 F.3d 844 (3d Cir. 1995); *United States v. Rutherford*, 104 F. Supp. 2d 1140 (D. Neb. 2000); *United States v. Hines*, 55 F. Supp. 2d 62 (D. Mass. 1999); *United States v. Starzecpyzel*, 880 F. Supp. 1027 (S.D.N.Y. 1995); *United States v. Santillan*, No. CR-96-40169 DLJ, 1999 WL 1201765 (N.D. Cal. Dec. 3, 1999); *United States v. Battle*, No. 98-3246, 1999 WL 596966 (10th Cir. Aug. 6, 1999); *United States v. Ruth*, 42 M.J. 730 (A. Ct. Crim. App. 1995); and a second opinion in *United States v. Ruth*, 46 M.J. 1 (C.A.A.F. 1997). I treat opinions on the standard legal databases as reported, even if some of them are technically to be treated as "unreported" under the rules of particular jurisdictions (which is not the case with *Santillan*). In the spirit of fullest disclosure, it should be noted that the author consulted in *Starzecpyzel*; that the author's friend and colleague Dr. Michael J. Saks was a defense witness in *Starzecpyzel*, *Rutherford*, and *Brown*; and that the author's friend and colleague Professor Mark P. Denbeaux was a defense witness or proffered witness in *Ruth*, *Velasquez*, *Paul*, and *Hines*.

53. 880 F. Supp. 1027 (S.D.N.Y. 1995).

54. *United States v. Starzecpyzel*, 880 F. Supp. 1027 (S.D.N.Y. 1995).

55. *Id.* at 1028. Much of the treatment of *Starzecpyzel* and *Velasquez* is drawn from D. Michael Risinger, *Handwriting Identification*, in 2 MODERN SCIENTIFIC EVIDENCE, *supra* note 13, § 22, at 79-123.

56. *Starzecpyzel*, 880 F. Supp. at 1028.

57. *Id.*

58. *Id.*

59. *See id.* at 1036.

Were the Court to apply *Daubert* to the proffered FDE testimony, it would have to be excluded. This conclusion derives from a straightforward analysis of the suggested *Daubert* factors – testability and known error rate, peer review and publication, and general acceptance – in light of the evidence adduced at the *Daubert* hearing.

*Id.*

experts are not practicing a science within the meaning of *Daubert*, *Daubert*'s validation requirements therefore do not apply.<sup>60</sup> (This approach obviously does not survive *Kumho Tire*.) He then, as already noted, analogizes such a proffered expert to a harbor pilot who learns how to do something dependably by experience.<sup>61</sup> As to whether the court would allow the prosecution's expert to testify to his conclusion that the signatures on the documents were forgeries based on his examination of numerous genuine signatures of the putative victim, the court says that the defense had:

presented no evidence, beyond the bald assertions [of its experts], that FDEs [forensic document examiners] cannot reliably perform this task. Defendants have simply challenged the FDE community to prove that this task can be done reliably. Such a demonstration of proof, which may be appropriate for a scientific expert witness, has never been imposed on "skilled" experts.<sup>62</sup>

Judge McKenna then declares himself persuaded that the inferences as to genuineness of the signature at issue in the case before him "can be performed with sufficient reliability to merit admission."<sup>63</sup>

Finally, it should be noted that (in anticipation of *Kumho Tire*) Judge McKenna emphasizes that he is dealing only with the limited issue of the skill of comparing a known signature with a questioned signature to determine whether the person whose name was reflected actually signed the questioned

---

60. *Id.*

61. *Id.* at 1029.

62. *Id.* at 1046. The implication that the ultimate risk of non-persuasion as to reliability is *ever* on the opponent of a proffer of evidence is startling, in light of Rule 104 of the Federal Rules of Evidence and the general notion that the party seeking admission must affirmatively convince the court of admissibility once the matter is seriously put in issue. Actually, Judge McKenna seems to have been aware of the problems that would be created by formally placing the burden of persuasion on the opponent of a proffer. The actual position taken by his opinion on that issue is ambiguous and unclear and, one must conclude, intentionally so. In the only explicit discussion of the issue, he concedes that Professor Berger takes the standard position that the burden is on the proponent of admissibility. *Id.* at 1031. He then cites an unexamined single line claim to the contrary from the middle of an article by a products liability practitioner whose main position is that *Daubert*'s effect should be viewed as allowing more things to be admitted, not less. *Id.* (citing Maskin, *supra* note 7, at 1936). However, Judge McKenna attempts not to choose between these two positions, characterizing the question before him as "legal" rather than factual, as if that makes the problem go away. *Id.* His later language from the passage cited in the text of this article seems to show his functional adoption of the problematical Maskin position; however, as the text indicates, he goes on to say that he is persuaded that handwriting identification testimony "can be performed" with "sufficient reliability to merit admission." *Id.* at 1046.

63. *Id.* Judge McKenna went on to fashion a jury instruction to be given in advance of the expert's testimony to explain that the testimony was not the result of a scientific process, so that the jurors would have no misconceptions in that regard. *Id.* at 1050-51.

signature. Judge McKenna makes clear that his analysis does not apply to any other asserted skill or global claim of expertise. Later courts and commentators, who have tended to treat *Starzeczyzel* as if it dealt with global validity, have generally missed this point.

In *United States v. Ruth*,<sup>64</sup> the Court of Military Appeals dealt with the issue of the reliability of forensic handwriting identification.<sup>65</sup> The court first notes that military courts have accepted handwriting expertise "for at least the past forty-four years."<sup>66</sup> It then declares that *Daubert* did not apply to nonscientific expertise, cites *Starzeczyzel* to establish the nonscientific nature of questioned document examination, and declares:

[I]t has been generally understood that expert testimony on handwriting comparison can assist panel members by focusing their attention on minute similarities and dissimilarities between exemplars that panel members might otherwise miss when they perform their own visual comparison . . . . It is largely in the location of these similarities and differences that a professional documents examiner has an advantage over panel members.<sup>67</sup>

As authority for this proposition, the court points only to an unpublished opinion<sup>68</sup> and concludes that the challenged handwriting identification testimony is admissible as helpful to the trier of fact under Rule 702. This unanalyzed global approach is now clearly unavailable after *Kumho Tire*.

It is important to note what application of expertise was being claimed reliable in *Ruth*. It was not the ability to determine if a signature was genuine, as was the case in *Starzeczyzel*. It was the much more questionable ability to attribute the authorship of a very small sample of writing (like a forged signature) to a particular person based on comparison to examples of the asserted forger's true writing.<sup>69</sup>

64. 42 M.J. 730 (A. Ct. Crim. App. 1995).

65. *United States v. Ruth*, 42 M.J. 730, 731 (A. Ct. Crim. App. 1995).

66. *Id.* at 732.

67. *Id.* at 733.

68. *See id.* (citing *United States v. Buck*, No. 84 Cr. 220-CSH, 1987 WL 19300 (S.D.N.Y. Oct. 28, 1997)).

69. Standard Osbornian theory of handwriting identification holds the latter a much easier task than the former. *See* Risinger & Saks, *supra* note 37, at 73 (discussing theories of handwriting identification). In this regard, consider the following quotations from three of the most respected authorities in the standard document examiner literature:

It is much easier to show that a fraudulent signature is not genuine than it is to show that such a writing is actually the work of a particular writer. . . . It would be strange indeed if in the few letters of a forged signature a forger would incorporate a sufficient number of the characteristics of his own writing actually to identify him.

ALBERT S. OSBORN, *QUESTIONED DOCUMENTS* 18 (2d ed. 1929).

There are two requirements which must be satisfied before a positive identification

The *Ruth* case involved a get-rich-quick scam that worked like this: Someone opened a bank account in Lichtenstein in the fictitious name "William Cooper" using a falsified copy of a passport.<sup>70</sup> The imposter then gained access to the personnel and pay records of approximately thirty-five American soldiers stationed at a base in Bamberg, Germany.<sup>71</sup> Using the information on bank accounts in those records, the impostor sent letters to the (American) banks of the soldiers directing wire transfers of the complete balance of their accounts to the "William Cooper" account at "one a.m. Eastern Time Zone on 01 May 1992."<sup>72</sup> Apparently, the letters were typed, with handwritten signatures. The scheme was uncovered when the depositors told their banks that the letters were fraudulent.<sup>73</sup> It is not completely clear whether the depositors notified their banks before or after transfers occurred, but it appears to have been before, as a result of bank inquiries concerning these unusual balance transfer directives.

Suspicion fell upon Private Joseph M. Durocher and Specialist Jeffrey A. Ruth, who were personnel action clerks in Bamberg with access to the relevant bank information.<sup>74</sup> Authorities interrogated Durocher, who apparently cooperated with prosecutors, confessed to the scheme, and implicated Ruth. Durocher's testimony was the main evidence against Ruth. The only corroboration of Durocher's story was a questioned document examiner's testimony that Ruth signed one of the forged signatures on letters to banks and that Ruth wrote one (but not all) of the signatures of William Cooper on the applications used to open the bank account in Lichtenstein.<sup>75</sup>

---

can be made. First of all, the forged signature must have been written in the natural handwriting of the forger. Signatures for the most part are short, and even a moderate degree of disguise may prevent accurate identification of a single specimen. . . . [If the questioned document is not in the natural hand of the forger] the likelihood of identification is very remote . . . . The entire problem is an extremely difficult one, and if not handled carefully, can lead to serious errors.

Ordway Hilton, *Can the Forger Be Identified from His Handwriting?*, 43 J. CRIM. L., CRIMINOLOGY & POLICE SCI. 547, 548, 555 (1953).

[W]hile it is often possible to express and justify a definite opinion as to whether a signature is genuine or forged, it is rarely that the identity of the forger can be established by comparing the handwriting of the forgery with specimens of the handwriting of suspects.

WILSON R. HARRISON, *SUSPECT DOCUMENTS: THEIR SCIENTIFIC EXAMINATION* 374 (2d ed. 1966).

70. *Ruth*, 42 M.J. at 731.

71. *Id.*

72. *Id.*

73. *Id.*

74. *United States v. Ruth*, 46 M.J. 1, 2 (C.A.A.F. 1997).

75. *Id.*

By now the reader will see the problem. In *Ruth*, the writing sample consisted of only fourteen to sixteen letters, which might or might not have represented an attempt to simulate the writing of the named signatory. Under a proper *Kumho Tire* approach, the issue would have been what, if anything, establishes that the questioned document examiner can reliably identify the writer of such a sample. However, the Court of Military Appeals in *Ruth* neither asks nor answers this question.

*United States v. Velasquez*<sup>76</sup> presents a similar problem.<sup>77</sup> The case deals with the conviction of Velasquez under 21 U.S.C. § 848, for engaging in a continual criminal enterprise involving five or more people.<sup>78</sup> The *only* evidence establishing that five people, rather than three, were involved in the alleged drug scheme was testimony by a government questioned documents examiner that two alleged co-participants of the defendant had written, at least in part, mailing labels used to ship drugs.<sup>79</sup> Thus, under a proper *Kumho Tire* approach, the issue would be whether and under what conditions questioned document examiners can reliably attribute authorship of individual parts of a document, the whole of which is extremely short, to particular individuals. Again, the *Velasquez* court neither asks nor answers this question. Instead, the court once again adopts, without analysis and by default, what functionally was a guild test. The *Velasquez* court's "reliability" analysis consists of merely reciting the document examiner's training, experience, and her own assertion that she had properly performed the analysis. The court declares that this established sufficient reliability.<sup>80</sup> Again, *Kumho Tire*'s emphasis on reliability of expertise in regard to the task to which it is applied in the particular case would seem to dispose of a guild test as a dispositive approach, especially when globally applied, as it was in *Velasquez*.

This brings us to *United States v. Jones*.<sup>81</sup> In many ways, the Sixth Circuit Court of Appeals's opinion in *Jones* is the standard against which all other unsatisfactory treatments of reliability issues in any area must be judged.

---

76. 64 F.3d 844 (3d Cir. 1995).

77. *United States v. Velasquez*, 64 F.3d 844 (3d Cir. 1995).

78. *Id.* at 845-46.

79. *Id.* at 846.

80. The *Velasquez* court initially takes the position that *Daubert* did not apply to handwriting because it was not science. *Id.* at 850. However, the court then claimed that "in an exercise of caution" it would review the proffered expertise "for qualifications, reliability, and fitness as those factors have been explicated in *Daubert*." *Id.* It then proceeds to never mention any reliability criteria beyond the testimony of the document examiner as to the very standard practice that was the subject of the challenge, and her experience. *Id.* at 850-51.

81. 107 F.3d 1147 (6th Cir. 1997).

In *Jones*, a thief obtained a credit card promotional mailing sent to Kathleen Jones's daughter's husband's aunt and uncle, on whose property the daughter and her husband lived in a house trailer.<sup>82</sup> The thief then rented a post office box in the name of a third party (who happened to be a co-worker of the defendant), filled out the credit card application, and requested that the issuer send the credit card to the post office box.<sup>83</sup> When the card arrived, the thief charged \$3748.00 worth of items over a two week period.<sup>84</sup> The credit card company's investigation implicated Kathleen Jones. Part of the evidence against Jones<sup>85</sup> was testimony by a questioned document examiner.<sup>86</sup> The Court of Appeals's opinion somewhat obscures the exact nature of the document examiner's testimony by stating that "Jones's signature was on: (1) the credit card application; (2) a post-office box registration form for the post-office box to which the card was sent; and (3) two Howard Johnson's motel registration forms, which contained the fraudulently procured Visa number at issue."<sup>87</sup> However, Jones's signature was on none of these documents. What was on the documents was what *purported* to be the signature of the aunt, which the questioned document examiner attributed to Jones. So the core *Kumho Tire* reliability issue in *Jones* is virtually the same as the one in *Ruth* and close to the one in *Velasquez*. As in those cases, the reliability issue is neither asked nor answered. However, the way in which it is not answered is what raises *Jones* to new heights of unsatisfactory judgecraft.

The *Jones* opinion first spends a page disposing of a trivial challenge to the authentication of an exemplar.<sup>88</sup> It then spends over five pages discussing the standard of review to be applied<sup>89</sup> (this was prior to the Supreme Court's

---

82. United States v. Jones, 107 F.3d 1147, 1149 (6th Cir. 1997).

83. *Id.*

84. *Id.*

85. There was not insignificant evidence against Jones independent of the handwriting identification. In addition to the coincidence that the post office box was in the name of a co-worker whose purse had been rifled when Jones was around, there was an identification of Jones by the motel clerk, and an identification of her by her own daughter from a bank camera photo (at a bank where she had no account) showing her doing a transaction at the approximate time the credit card was used to obtain money from that bank. Indeed, one significant issue, had the courts bothered to go into the issues properly, would have been whether the document examiner was privy to such information and its conscious or unconscious effect on his conclusion. See Risinger & Saks, *supra* note 37, at 64 and accompanying notes (noting that handwriting analysis is influenced by contextual cues).

86. *Jones*, 107 F.3d at 1149.

87. *Id.*

88. *Id.* at 1149-50.

89. *Id.* at 1150-56.

opinion in *General Electric Co. v. Joiner*<sup>90</sup>). It then spends two and a half pages deciding that *Starzecpyzel* was right, that handwriting expertise is not science, and that *Daubert* is therefore irrelevant<sup>91</sup> (this was prior to *Kumho Tire*, of course). The court then says "without relying on *Daubert*, we now address whether handwriting analysis constitutes 'technical, or other specialized knowledge' under the Federal Rules of Evidence and whether the expert handwriting analysis offered in this case was sufficiently reliable."<sup>92</sup> The court spends less than two and a half pages on these core issues of the case.<sup>93</sup> In this sparse treatment, the court initially makes two points it appears to think are persuasive on issues of general reliability: (1) Other courts and commentators have uniformly found or assumed that handwriting identification expertise is (globally) a proper subject of court testimony and that the appellant is therefore "asking us to do what no other court we have found has done,"<sup>94</sup> and (2) "the Federal Rules of Evidence themselves suggest that handwriting analysis is a field of expertise."<sup>95</sup> In trying to justify this latter statement, the court sets out an egregious version of what I call "the Rule 901(b)(3) fallacy,"<sup>96</sup> and flagrantly misquotes the Rule in order to do it.

Rule 901(a) sets out the general standard for authentication of any evidence whatsoever.<sup>97</sup> Rule 901(b) gives a non-exhaustive list of acceptable recurrent means of satisfying the general requirements of Rule 901(a) for "evidence sufficient to support a finding that the matter in question is what its proponent claims."<sup>98</sup> One way of doing this is set out in Rule 901(b)(3): "Comparison by the trier of fact or by expert witnesses with specimens which have been authenticated."<sup>99</sup> This is the full text of the rule. It is not limited to documents (much less handwriting), nor does it reference them. In common practice it has been applied whenever an area of expertise, from fingerprints to DNA analysis, is shown or conceded to be reliable to make such comparisons

---

90. See *General Elec. Co. v. Joiner*, 522 U.S. 136, 139 (1997) (setting "abuse of discretion" as proper standard by which to review district court's decisions to admit or exclude expert scientific evidence).

91. *United States v. Jones*, 107 F.3d 1147, 1156-59 (6th Cir. 1997).

92. *Id.* at 1159.

93. *Id.* at 1159-61.

94. *Id.* at 1159.

95. *Id.*

96. FED. R. EVID. 901(b)(3).

97. FED. R. EVID. 901(a).

98. FED. R. EVID. 901(b). Presumably, by implication, is what its proponent claims *that renders the evidence relevant*, under the standards of Rule 401.

99. FED. R. EVID. 901(b)(3).

under the standards of Rule 702.<sup>100</sup> However, Rule 901(b)(3) most certainly does not contain any suggestion that because there is a claim that an expert is comparing specimens, the existence of Rule 901(b)(3) means that he or she meets the reliability requirements of Rule 702. This is flagrantly backward reasoning, whether applied to handwriting identification expertise or anything else. Pursuant to *Daubert* and *Kumho Tire*, the court must determine the existence of reliable expertise under Rule 702 before it can consider the evidence at all. Only if reliable expertise exists under the standards of Rule 702 may the court reference Rule 901(b)(3) to establish sufficiency for authentication purposes (assuming the expertise involves a comparison of standards.)

The mistaken argument from Rule 901(b)(3) just given is the standard form of the fallacy often put forward in prosecution briefs.<sup>101</sup> However, the *Jones* court takes it one step further (presumably as a result of an embarrassing failure to read the actual text of the rule) when it says that Rule 901(b)(3) "provides for authentication of a document by [c]omparison by . . . expert witnesses with specimens which have been authenticated."<sup>102</sup> The court then compounds its error by claiming that if handwriting identification expertise were excluded "there would be no place for expert witnesses to compare writing on one document with that on another in order to authenticate a document. In other words, appellant's approach would render Rule 901(b)(3) meaningless."<sup>103</sup> This might be true if the rule said what the court apparently thinks it says, but plainly, it does not.

To its credit, the court realizes, in instinctive anticipation of *Kumho Tire*, that what it has written does not "guarantee the reliability or admissibility of this type of testimony in a particular case. Because this is non-scientific testimony, its reliability largely depends on the facts of each case."<sup>104</sup> However, it then sets out an approach to "reliability" that barely deals with actual reliability. Perhaps not surprisingly, it adopts a global combination of the experience test and the guild test.

*Jones*, as already noted, was not the first court to adopt the guild test. However, the unintentional humor of the opinion shows forth most strongly in the details that the *Jones* court seems to believe help to justify an inference

---

100. See WRIGHT & GOLD, FEDERAL PRACTICE AND PROCEDURE, EVIDENCE § 7108 n.2-15 (collecting cases applying Rule 901(b)(3)).

101. See, e.g., Pre-Trial Transcript at 15-17, *United States v. McVeigh*, No. 96-CR-68, 1997 WL 47724 (D. Colo. Feb. 5, 1997) (rejecting government's attempt to use argument based on mistaken reading of Rule 901(b)(3)).

102. *United States v. Jones*, 107 F.3d 1147, 1159 (6th Cir. 1997).

103. *Id.* (emphasis added).

104. *Id.* at 1160.

of reliability. After describing a fairly normal training history for a government questioned document examiner, the court notes that his primary job responsibilities consisted of the "examination and comparison of questioned handwriting."<sup>105</sup> The court then notes that the witness

estimated that throughout his career, he had conducted "well over a million comparative examinations." In addition he has published numerous articles in the field and testified approximately 240 times in various courts. To put it bluntly, the federal government pays him to analyze documents, the precise task he was called upon to do in the district court.<sup>106</sup>

This may be the only case on record in which a person's government job description has been taken as evidence of the reliability of asserted expertise. Even more amusingly, the passage illustrates the credulity of the court and the collision between the court's approach and any mildly skeptical approach to the dependability of information. This witness testified to conducting "well over a million comparative examinations."<sup>107</sup> If he had been doing document examination eighteen hours a day, every day for fifty years, he still would have to have done more than three comparative examinations per hour to reach a million.<sup>108</sup> Yet the court accepts the testimony without objection and cites it as substantiation for the reliability of his expertise. The *Jones* court then cites Professor Imwinkelried for the proposition that the reliability of non-scientific expert testimony increases with the more experiences an expert has had and the similarity of those experiences to the expert's testimony.<sup>109</sup> However, the cited article makes no such sweeping statement at the cited pages or anywhere else. This is not surprising, because Professor Imwinkelried does not generally make unsophisticated global statements. Imwinkelried's article does take the position that some kinds of inferences must be

---

105. *Id.*

106. *Id.* (citations omitted).

107. *Id.* at 1161. Lest the reader believe that some transcription error has been made, the trial transcript does indeed reveal this to have been Mr. Sperry's testimony. Transcript of Testimony, at 123, *United States v. Jones*, No. CR 3-95-24 (E.D. Tenn. June 29, 1995) (on file with author).

108. The real circumstances are even more extreme. Mr. Sperry began his training (a two year course) in 1979. Transcript of Testimony, at 121, *United States v. Jones*, No. CR 3-95-24 (E.D. Tenn. June 29, 1995) (on file with author). Thus, at the time of trial, June 1995, he had 16 years of experience, 14 of which were post-training. This more than triples his actual hourly output, to over nine for every waking hour. Yet, besides his "well over a million comparative examinations," he also testified to having been assigned to "7300 [or] 7400 cases." *Id.* at 123. This works out to at least one and a quarter cases, every day including Sundays and holidays, without a break, for all 16 years, including his training period.

109. *United States v. Jones*, 107 F.3d 1147, 1160 (6th Cir. 1997) (citing Imwinkelried, *supra* note 13, at 2292-93).

based on much experience as a precondition to any claim to reliability, but it never says that such experiences alone guarantee, or even necessarily increase, accuracy in all cases.<sup>110</sup>

The court then continues its unaccountable course by asserting that "handwriting examiners themselves have recognized the importance of experience"<sup>111</sup> (no doubt true), but it supports this assertion with a quote from an article in the *Journal of Forensic Document Examination* that actually claims that the bulk of document examiner experience with handwriting forms is obtained outside the professional sphere and is common with the rest of the world. "For handwriting examiners, this experience comes mainly from the exposure we have to handwriting throughout the course of our life, the majority of which normally would occur before specializing in forensic handwriting examination."<sup>112</sup>

It is on these grounds, coupled with the fact that the document examiner described the process by which he arrived at his conclusions, that the court declares "given Sperry's various training experiences, his job responsibilities, his years of practical experience, and the detailed nature of his testimony in this case, the court did not abuse its discretion by admitting his testimony."<sup>113</sup>

If the *Jones* court's careless handling of both sources and reasoning is unusual, *United States v. Paul*<sup>114</sup> is in some ways stranger still.<sup>115</sup> In *Paul*, the Court of Appeals's statement of the facts and history of the case is precise and pertinent, and will be set out here in its entirety.

#### I. FACTS

In May 1996, an unidentified person who stated that he was a bank investigator telephoned Ed Spearman, branch manager of Wachovia National Bank (Wachovia) at Atlanta, Georgia, and warned him that someone intended to leave a note at the bank in an attempt to extort money from the bank. The "investigator" instructed Spearman to follow the directions in the note. Spearman contacted bank security and the Federal Bureau of Investigation (FBI), who advised him to contact the agency immediately if he received an extortion demand. On the following morning, a security camera outside the entrance to Wachovia Bank videotaped a man, wearing a scarf and sunglasses, place an envelope under the front door of the bank. Inside the envelope, addressed to Spearman, was an extortion note that

---

110. Imwinkelried, *supra* note 13, at 2292-94.

111. *Jones*, 107 F.3d at 1160.

112. *Id.* (quoting Bryan Found & Doug Rogers, *Contemporary Issues in Forensic Handwriting Examination: A Discussion of Key Issues in the Wake of the Starzecpyzel Decision*, 8 J. FORENSIC DOCUMENT EXAMINATION 1, 26 (1995)).

113. *Id.* at 1161.

114. 175 F.3d 906 (11th Cir. 1999).

115. *United States v. Paul*, 175 F.3d 906 (11th Cir. 1999).

directed Spearman to deliver \$100,000 to the men's restroom of a downtown Atlanta McDonald's restaurant. The note threatened violence if Spearman did not follow the instructions and make the payment. Spearman notified bank security and the FBI.

The investigating agents developed a plan to arrest the extortionist: an FBI agent, acting as Spearman, would drive Spearman's car to the McDonald's and place a briefcase in the men's restroom, while surveillance agents would watch the restroom and arrest the person who took the briefcase.

In executing the plan, FBI Agent Eric Bryant testified that upon his arrival at the McDonald's, he entered the men's restroom, observed appellant Sunonda Paul in a restroom stall, left a briefcase and exited the restroom. FBI surveillance agents testified that they later saw Paul sitting at a table near the restroom. As Bryant left the McDonald's, surveillance agents observed Paul enter the restroom again and then attempt to leave the establishment with the briefcase in his backpack. When confronted, Paul told the agents that he was in the area to visit a nearby gym and had stopped at the McDonald's for breakfast. He also told them that he decided to take the briefcase after he found it in the restroom. Paul, however, was dressed in casual street clothing and had no gym clothes or athletic equipment in his possession. The agents arrested him.

## II. PROCEDURAL HISTORY

A grand jury indicted Paul on one count of bank extortion, in violation of 18 U.S.C. § 2113(a), and Paul pleaded not guilty. Prior to trial, Paul moved in limine to exclude FBI document examiner Larry Ziegler's testimony regarding handwriting analysis. The district court, however, denied Paul's motion at the pretrial hearing.

The demand note left at Wachovia was the key evidence in determining whether Paul was the extortionist. Although FBI agents examined the videotape to determine the identity of the person who delivered the note, they could not identify the person conclusively. Consequently, the FBI conducted fingerprint and handwriting analysis tests on the note to establish the identity of the extortionist. A fingerprint expert concluded that the latent prints on the note and envelope did not match Paul's fingerprints.

Ziegler, the FBI document examiner, compared the handwriting on the note and the envelope to Paul's handwriting samples and concluded that Paul was the author of both. Specifically, Ziegler asked Paul to write the word restaurant. In the presence of an FBI agent, Paul misspelled the word as follows: "resturant." In the extortion note the extortionist misspelled the word restaurant the same way. Ziegler also asked Paul to write out "Spearman." Paul spelled it "Sperman," the same way the extortionist had addressed the envelope.<sup>116</sup>

---

116. *Id.* at 908-09.

Oddly, the court seems to turn the case into one concerning the dependability of reasoning about the authorship of a document from misspellings in exemplars, without realizing that this is not necessarily an expertise issue. Additionally, the practice is somewhat controversial even in document examiner literature, and it has little to do with the reliability of assignment of authorship based on comparison of form.

Suppose an investigator with no claimed skill in document examination notices what appears to him to be unusual misspellings in a typed robbery note (the same misspelling, say, that was in the robbery note in Woody Allen's movie *Take the Money and Run*, which said (in part), "I am pointing a gub at you"). Acting on other information, he obtains a search warrant for the residence of a suspect and discovers numerous documents, typed or not, in which the suspect refers to "gubs" in contexts which clearly indicate he meant guns ("the NRA is right to oppose gub control legislation"). Virtually every court would receive such evidence authenticated by the investigator, though he claims no special knowledge about the uncommonness of this particular misspelling of "gun." Such a case raises interesting issues of jury notice and the accuracy of jury notice in regard to base rate occurrences of misspellings derived from common experience, but it does not raise issues of *Daubert/Kumho Tire* reliability of expertise. Exactly how much a document examiner ought to rely on or be influenced by misspellings is a subject of some controversy, and document examiner literature contains warnings against assuming uncommonness and making too much of misspellings.<sup>117</sup> All of this appears to have escaped the notice of the Court of Appeals, however, because the opinion never mentions these issues.

Of course, though it is not explicitly noted by the court, the questioned document examiner in the case did perform a comparison-of-form analysis in addition to noting the misspellings, and it was that identification by comparison of form that was the subject of Paul's reliability objection. The court deals with that objection as follows:

Paul argues that Ziegler's testimony is not admissible under the *Daubert* guidelines because handwriting analysis does not qualify as reliable scientific evidence. His argument is without merit. In *Daubert*, the Supreme Court held that Federal Rule of Evidence 702 controls decisions regarding the admissibility of expert testimony. The Supreme Court declared that under Rule 702, when "[f]aced with a proffer of expert scientific testimony[,] the trial judge must determine at the outset pursuant to Rule 104(a), whether the expert is proposing to testify to (1) scientific knowledge that (2) will assist the trier of fact to understand or determine a

---

117. See Risinger et al., *supra* note 3, at 770-71 (citing warning against mistaken interpretations of misspellings).

fact in issue." The Supreme Court stated that "[t]he inquiry envisioned by Rule 702 is, we emphasize, a flexible one" and that "Rule 702 . . . assign[s] to the trial judge the task of ensuring that an expert's testimony both rests on a reliable foundation and is relevant to the task at hand." The Court also listed several factors to assist in the determination of whether evidence is scientifically reliable. . . .

Many circuits were split at the time of trial, however, on whether *Daubert* should apply to nonscientific expert testimony. Some held that the application of *Daubert* is limited to scientific testimony, while others used *Daubert's* guidance to ensure the reliability of all expert testimony presented at trial.

Recently, however, in *Kumho Tire Company, Ltd. v. Carmichael*, the Supreme Court held that *Daubert's* "gatekeeping" obligation, requiring the trial judge's inquiry into both the expert's relevance and reliability, applies not only to testimony based on "scientific" testimony, but to all expert testimony. The Court further noted that Rules 702 and 703 give all expert witnesses testimonial leeway unavailable to other witnesses on the presumption that the expert's opinion "will have a reliable basis in the knowledge and experience of his discipline." Moreover, the Court held that a trial judge may consider one or more of the specific *Daubert* factors when doing so will help determine that expert's reliability. But, as the Court stated in *Daubert*, the test of reliability is a "flexible" one, and *Daubert's* list of specific factors neither necessarily nor solely applies to all experts or in every case. Alternatively, *Kumho* declares that "the law grants a district court the same broad latitude when it decides how to determine reliability as it enjoys in respect to its ultimate reliability determination."<sup>118</sup>

And that is all, folks. The careful (or even careless) reader will have noted that, beyond declaring in the second line of the passage that Paul's "argument is without merit," the court never addresses the reliability issue, nor does it address what, if anything, was before the district court that would not have rendered its determination of reliability an abuse of discretion. There is no formulation of "the task at hand," no description of a reliability test, no reference to information before the district court or the district court's reasoning, nothing.

Having assumed the conclusion of sufficient reliability with no analysis of the issue whatsoever, the rest of the opinion on admissibility follows as a matter of course, finding after a recitation of the expert's credentials that the testimony of a qualified expert could assist the trier of fact and that its prejudicial effect did not substantially outweigh its probative value.<sup>119</sup> All of this is based on the assumed conclusion to the reliability issue that the court

---

118. *Paul*, 175 F.3d at 909-10 (footnotes omitted) (citations omitted).

119. *Id.* at 911.

never explicitly addresses, but given the recitative as to credentials, what emerges in the end is functionally the guild test.

To complete the Court of Appeals's reliability opinions extant as of this writing, we turn to the Tenth Circuit Court of Appeals's opinion in *United States v. Battle*.<sup>120</sup> Battle was accused of coming from New York to Kansas to set up a drug distribution operation.<sup>121</sup> The evidence against him was voluminous and involved many witnesses and many episodes.<sup>122</sup> The handwriting identification testimony figured in a single episode not particularly central to the case, but nevertheless relevant because, if believed, it would establish that Battle received money surreptitiously from an out-of-state source under suspicious circumstances. Western Union had a record of a money transfer showing a "Tyler Evans" as the sender and "Anthony Jenkins" as the receiver.<sup>123</sup> The questioned document examiner was called to testify that he had examined exemplars of Battle's signature given when he was booked and, in his opinion, Battle signed the name "Anthony Jenkins" to the money transfer.<sup>124</sup>

As usual, the court fails to realize that the document examiner was testifying to one of the sub-tasks in handwriting comparison most likely to be unreliable. There are eleven letters in "Shawn Battle" and fourteen in "Anthony Jenkins" (which was signed only once). The two names share no capital letters, no letter combinations, and only four small letters (e, h, n, and t). One sample is in a presumably normal signature hand. The other is not, unless someone named "Anthony Jenkins" in reality signed his own name. If the Tenth Circuit had done what the Supreme Court did in *Kumho Tire*, it would have examined the reasons to believe or to doubt the accuracy of such a claimed identification. Although it does cite *Kumho Tire* in a pro forma way, it does no such analysis. It merely recites the document examiner's credentials (the guild test again) and declares that "[o]ur study of the record on appeal convinces us that McPhail's proffered testimony met the reliability and relevancy test of *Daubert*."<sup>125</sup> The opinion manifests some discomfort about its own conclusion, however, as it goes on to say "[b]e that as it may, in any event any error in this regard is, in our view, harmless error when the evidence is considered as a whole."<sup>126</sup>

---

120. No. 98-3246, 1999 WL 596966 (10th Cir. Aug. 6, 1999).

121. *United States v. Battle*, No. 98-3246, 1999 WL 596966, at \*1 (10th Cir. Aug. 6, 1999).

122. *Id.* at \*1-2.

123. *Id.* at \*3.

124. *Id.*

125. *Id.* at \*4.

126. *Id.*

If the district court opinion in *Starzecpyzel* started the judicial struggle with handwriting identification reliability, the opinion in *United States v. Hines*<sup>127</sup> contrasts sharply with the run of Court of Appeals opinions in taking the reliability issue seriously.<sup>128</sup> On January 27, 1997, someone robbed the Broadway National Bank in Chelsea, Massachusetts, by using a demand or "stick up" note and escaped.<sup>129</sup> The teller who was robbed, Ms. Jeanne Dunne, described the perpetrator as a dark skinned black man with a wide nose and medium build.<sup>130</sup> Ms. Dunne is white, and the court characterized the description as "as close to a generic identification of an African American man as one can imagine."<sup>131</sup> Later, Dunne failed to pick Hines out of a mug-book where his picture appeared and failed to positively identify him from an eight picture photo spread, though she said Hines "resembled" the robber. Months later, however, she picked Hines out of a lineup and positively identified him at trial.<sup>132</sup>

The main corroboration of this eyewitness identification came from an FBI questioned document examiner who compared the robbery note with exemplars of Hines's handwriting and concluded that Hines had written the note. A trial on this evidence ended in a hung jury. Before the retrial, Hines moved to disallow the document examiner's testimony based on lack of sufficient reason to find it reliable under *Daubert*.<sup>133</sup> The court (Judge Gertner) granted the motion in part and wrote the published opinion during and after the second trial that also resulted in a hung jury, in order to explain its ruling and give guidance to the parties in the event of a third trial.<sup>134</sup>

As previously noted in regard to *Daubert*, Judge Gertner identifies what she takes to be a "mixed message" in both *Daubert* and in *Kumho Tire*, with the emphasis on reliability pointing in the direction of more rigor in the evaluation of expertise under Rule 702 and the emphasis on "the uniqueness of the trial setting, the 'assist the trier' standard and flexibility" doing the opposite.<sup>135</sup> Nevertheless, the court concludes that the main emphasis is on insuring sufficient reliability and that the Supreme Court "is plainly inviting

---

127. 55 F. Supp. 2d 62 (D. Mass. 1999).

128. *United States v. Hines*, 55 F. Supp. 2d 62 (D. Mass. 1999).

129. *Id.* at 63.

130. *Id.* at 71.

131. *Id.*

132. *Id.*

133. *Id.* at 64.

134. *Id.*

135. *Id.* at 65.

a reexamination even of "generally accepted" venerable, technical fields" such as handwriting identification."<sup>136</sup>

Judge Gertner accepts this invitation, with a number of caveats. Based on the hearings, she seems inclined to bar the questioned document examiner testimony in its entirety. However, "[t]his handwriting challenge was raised at the eleventh hour. The hearing was necessarily constrained by the demands of the imminent trial and the schedules of the experts. The court was unwilling on this record to throw out decades of "generally accepted" testimony."<sup>137</sup>

In addition, the "compromise solution" Judge Gertner accepts is "derived largely from the case law that pre-dated *Kumho*."<sup>138</sup> In other words, Judge Gertner is not entirely comfortable that she had fully digested and applied the broader implications of *Kumho Tire* that might have been inconsistent with this compromise. At any rate, Judge Gertner proceeds to distinguish between a questioned document examiner's testimony comparing the robbery note with the exemplars and identifying similarities and differences, and testimony concerning the document examiner's inferences of authorship based on those similarities. In sum, she allows the former and bars the latter, essentially adopting the similar approach of Judge Matsch in *United States v. McVeigh*,<sup>139</sup> whom she quotes at length.<sup>140</sup>

There is a certain commonsense appeal to this approach. The document examiner's extensive experience looking at handwriting may have sensitized the expert to the perception and identification of similarities or differences that an ordinary person might not notice. At any rate, the document examiner will be free to spend more time isolating such similarities and differences than we could expect jurors to do pursuant to their own examinations during deliberations. Viewed this way, document examiners appear to become summarizational witnesses, and the notion of "expertise" becomes much less central to their function.

However, there is a serious problem with this, especially if the document examiner is allowed to recite her credentials, titles, and job descriptions. By identifying a similarity or difference, the examiner is inevitably perceived as asserting the *significance* of those similarities or differences in regard to assigning authorship, so that juries can easily infer the barred conclusions. In

---

136. *Id.* at 67.

137. *Id.* at 67.

138. *Id.* at 68.

139. No. 96-CR-68, 1997 WL 47724 (D. Colo. Trans. Feb. 5, 1997).

140. *Hines*, 55 F. Supp. 2d at 70 (citing *United States v. McVeigh*, No. 96-CR-68, 1997 WL 47724, at \*\*3-4 (D. Colo. Trans. Feb. 5, 1997)). It should be noted that the author's friend and colleague Mark P. Denbeaux was the defense expert in *McVeigh* as well as *Hines*.

practice, this is profoundly true, because document examiners who believe they have identified the author of a writing by comparison will normally *only* point out similarities. If differences are called to their attention, they will dismiss them as not being significant or "real" differences, but merely manifestations of "individual variation." Nevertheless, the Solomonic compromise of Judges Matsch and Gertner is clearly an improvement over surrendering the gatekeeping function entirely to the guild, as most other courts have done. *United States v. Santillan*<sup>141</sup> suggests that the compromise is becoming common practice.<sup>142</sup>

In *Santillan*, the defendant, Rogelio Santillan, was charged with conspiracy to distribute false immigration documents.<sup>143</sup> According to the opinion of the court, the prosecution's document examiner proposed to testify "that she has identified, using control samples of defendant's handwriting, Santillan's handwriting on numerous "questioned" documents."<sup>144</sup> Note that in an opinion referring to *Kumho Tire*,<sup>145</sup> more than seven months after the Supreme Court's decision in *Kumho Tire*, this is as much information as we are given on the task at issue in the case before the court. Clearly, the requirements of *Kumho Tire* are not yet dependably appearing on the radar screens of the lower courts.<sup>146</sup> In any event, Judge Jensen then goes on to deal with the reliability issue globally. After briefly reviewing (and criticizing)<sup>147</sup> the extent research data, the court adopts the *Hines/McVeigh* approach.<sup>148</sup>

A similar result was reached in *United States v. Rutherford*.<sup>149</sup> In *Rutherford*, someone registered at a cattle auction as one "George Hipke," an actual person of repute.<sup>150</sup> The imposter then purchased an expensive lot of cattle and paid for them with a check bearing the name signature "George

---

141. No. CR-96-40169 DLJ, 1999 WL 1201765 (N.D. Cal. Dec. 3, 1999).

142. *United States v. Santillan*, No. CR-96-40169 DLJ, 1999 WL 1201765 (N.D. Cal. Dec. 3, 1999).

143. *Id.* at \*1.

144. *Id.*

145. *Id.*

146. The opinions in *Paul, Battle*, and *Hines* all were published after *Kumho Tire*, but sufficiently close in time that their failure to properly digest and apply it is perhaps understandable.

147. *Santillan*, 1999 WL 1201765, at \*3. Judge Jensen seemed especially troubled that Dr. Moishe Kam, the government's chief researcher on handwriting expertise issues, refused to share his raw data.

148. *Id.* at \*4.

149. 104 F. Supp. 2d 1190 (D. Neb. 2000).

150. *United States v. Rutherford*, 104 F. Supp. 2d 1190, 1191 (D. Neb. 2000). The details of the case are derived from the court's opinion, supplemented by a 9/21/00 telephone conversation with Adam Sipple, Esq. of the Clarence E. Mock Law Firm, Rutherford's attorneys.

Hipke." He then removed the cattle. The crime came to light when the real George Hipke learned of the check from his bank.

Rutherford was a retired banker in the region. An informant indicated that Rutherford had masterminded the scheme, although the people at the sale said that Rutherford was not the person who tendered the check. However, Rutherford was indicted for bank fraud. The proposed expert testimony was that Rutherford in fact signed a "buyer registration form" in Hipke's name, that Hipke did not sign the check bearing his name (a fact that was undisputed), that there was a high probability that Rutherford did sign it, and that inscriptions on the bottom of a "load out" sheet from the Columbus Sale Barn<sup>151</sup> were probably written by Rutherford.

Once again, although the court cites *Kumho Tire*, there is no further attempt to formulate the separate tasks involved in this scenario or to examine them individually. This is perhaps a bit surprising, given the extensive nature of the hearing that was undertaken. And once again, the main task at issue is the attribution of authorship of a forged signature based on the limited amount of writing in the signature, though in this case there was more than one signature to work with.

In any event, after criticizing the highly suggestive way in which the government presented the problem to the expert in obtaining his original opinion,<sup>152</sup> the court (Judge Bataillon) ultimately adopts the *Hines/McVeigh* approach, allowing the proffered expert to testify only by pointing out similarities and differences and not allowing any explicit opinions concerning authorship or probability of authorship.<sup>153</sup>

Finally, for the sake of completeness, we note *United States v. Brown*,<sup>154</sup> an unreported check forgery case in which, once again, the defendant disputed the reliability of the government's handwriting expert's testimony that the defendant authored the forged signature. After holding a *Daubert/Kumho* hearing, the court (Judge Collins), adopts the *Hines/McVeigh* approach by permitting the proffered expert to testify without rendering "an ultimate conclusion on who penned the questioned writings."<sup>155</sup>

It is interesting to note that in this set of eleven federal cases (including *McVeigh* and *Brown*), only seven of the eleven trial judges involved held full-

---

151. The only handwritten material besides signatures was the inscription "L & C Livestock" on the bottom of the load out sheet.

152. *Rutherford*, 104 F. Supp. 2d at 1193.

153. Thereafter, Rutherford was tried and acquitted in May of 2000.

154. No. CR 99-184 ABC (C.D. Cal. Dec. 1, 1999).

155. *United States v. Brown*, No. CR 99-184 ABC, slip op. at 9 n.6 (C.D. Cal. Dec. 1, 1999) (on file with author).

scale reliability hearings.<sup>156</sup> Of those who held such hearings, six out of seven manifested substantial reservations about handwriting identification expertise even on a global level.<sup>157</sup> Contrast this with the Court of Appeals judges who heard these cases, all of whom were institutionally insulated from *having* to come to grips with the actual state of knowledge concerning handwriting identification by cold records and appendices, and *all* of whom felt comfortable brushing off reliability challenges with some version of the guild test. The inertia of a venerable tradition of admissibility appears to be a powerful incentive to turn a blind eye to the evidence. In light of *Kumho Tire*, perhaps even appellate judges will have to pay attention to the issues raised by these cases as well as others involving non-science forensic science of questionable reliability.

#### *Postscript*

As this Article was being finalized, a decision was rendered in *United States v. Fujii*.<sup>158</sup> In this case, Judge Gottschall delivers the first decision in a handwriting identification case in modern times that entirely excludes proffered document examiner testimony on reliability grounds.<sup>159</sup> It is also the first opinion in a handwriting case to manifest a proper *Kumho Tire* "task at hand" approach to the reliability question.

In *Fujii*, it was alleged that defendant Masao Fujii had been involved in a scheme to obtain the fraudulent entry into the United States of two Chinese nationals. The Chinese nationals tendered certain hand-printed immigration forms in connection with their attempted entry at John F. Kennedy airport in New York in December of 1999. As evidence of Fujii's participation, the prosecution sought to call an Immigration and Naturalization Service document examiner who would testify that she had compared the printing on

156. *Ruth, Velasquez, Jones, and Battle* were disposed of at trial, either without a hearing or on voir dire of the government's proposed expert.

157. Judges McKenna, Matsch, Gertner, Jensen, Collins, and Bataillon were those six. Only Judge Tidwell, the trial judge in *Ruth*, seemed unconcerned. Apropos of this, considered the following quotation from Judge McKenna's opinion in *Starzecpyzel*:

If forensic document examination does rely on an underlying principle, logic dictates that the principle must embody the notion that inter-writer differences, even when intentionally suppressed, can be distinguished from natural variation. How FDEs might accomplish this was unclear to the Court before the hearing, and largely remains so after the hearing.

*United States v. Starzecpyzel*, 880 F. Supp. 1027, 1032 (S.D.N.Y. 1995). [As the Postscript indicates, Judge Gottschall must now be added to the list of judges with reservations, making it seven out of eight.]

158. No. 00 CR 17 (N.D. Ill., Sept. 25, 2000) (slip opinion on file with author).

159. *United States v. Fujii*, No. 00 CR 17, slip op. at 3-4 (N.D. Ill. Sept. 25, 2000).

the forms with examples of printing by Fujii, and that in her opinion Fujii printed the fraudulent forms. The defense objected on *Daubert/Kumho* grounds, and a hearing was held on the issue.

In her opinion, Judge Gottschall notes that in general "[h]andwriting analysis does not stand up well under the *Daubert* standards."<sup>160</sup> However, as to general issues of reliability, she concludes that "[t]his court need not weigh in on this question, . . . for whether handwriting analysis *per se* meets the *Daubert* standards, its application to this case poses more significant problems."<sup>161</sup> This is for two reasons. First, virtually all data on document examiner dependability in identifying the author of handwriting has dealt with cursive and not printed writing.<sup>162</sup> The single recorded proficiency test involving hand printing<sup>163</sup> revealed a 45% error rate on that test.<sup>164</sup> Second, and perhaps more importantly, Fujii is a native Japanese who learned to print in English as a second language in Japan. The defense submitted an affidavit of Mark Litwicki, Director of Loyola University's English as a Second Language program, who had substantial experience with teaching English to Japanese students in the United States and Japan.<sup>165</sup> The essence of his testimony was that Japanese students learn to print in English only after years of training in the exact copying of Japanese characters, in which uniformity is "an important and valued principle."<sup>166</sup> Because Japanese students "spend many years attempting to maximize the uniformity of their writing," the emphasis on uniformity carries over into their English writing.<sup>167</sup> Consequently, Litwicki testified, "it would be very difficult for an individual not familiar with the English handwriting of Japanese writers to identify the subtle dissimilarities in the handwriting of individual writers."<sup>168</sup>

Considering all this, the court concludes:

Does Ms. Cox [the document examiner] have any expertise which would allow her to distinguish between unique characteristics of an individual Japanese handprinter and characteristics that might be common to many or all native Japanese handprinters? In an analysis that depends entirely on

---

160. *Id.* at 3.

161. *Id.*

162. *Id.* at 4.

163. In *Fujii*, Dr. Michael J. Saks, as one of the witnesses for the defense, testified about the results of the 1986 Forensic Science Foundation test. The results of this test are reported in Risinger et al., *supra* note 3, at 746-47.

164. *Fujii*, slip op. at 4.

165. *Id.*

166. *Id.* at 5 (quoting witness Mark Litwicki).

167. *Id.*

168. *Id.*

what is similar between writing specimens and what is different, it would seem to this court essential that an expert have some ability to screen out characteristics which might appear eccentric to the writer, compared to native English printers, but which might in fact be characteristic of most or all native Japanese writers, schooled in English printing in Japan, in printing English. There is no evidence in the record that Ms. Cox has such expertise or has even considered the problem Mr. Litwicki has pointed out.

Considering the questions about handwriting analysis generally under *Daubert*, the lack of any evidence that the identification of handprinting is an expertise that meets the *Daubert* standards and the questions that have been raised – which the government has not attempted to answer – about its expert's ability to opine reliably on handprinting identification in dealing with native Japanese writers taught English printing in Japan, the court grants the defendant's motion [to exclude].<sup>169</sup>

Judge Gottschall's opinion is a masterful example of particularized "task at hand" analysis under the standards of *Kumho Tire*. It provides a model for other courts in how to approach the reliability of specific applications of all forms of non-science forensic science.

---

169. *Id.* at 5-6.