Tulsa Law Review

Volume 43
Issue 2 *Daubert, Innocence, and the Future of Forensic Science*

Winter 2007

Goodbye to All That, or a Fool's Errand, By One of the Fools: How I Stopped Worrying about Court Responses to Handwriting Identification (and Forensic Science in General) and Learned to Love Misinterpretations of Kumho Tire v. Carmichael

Michael D. Risinger

GOODBYE TO ALL THAT, OR A FOOL'S ERRAND, BY ONE OF THE FOOLS:

HOW I STOPPED WORRYING ABOUT COURT RESPONSES TO HANDWRITING IDENTIFICATION (AND "FORENSIC SCIENCE" IN GENERAL) AND LEARNED TO LOVE MISINTERPRETATIONS OF *KUMHO TIRE V. CARMICHAEL*[*]

D. Michael Risinger[**]

"[H]andwriting is even more precise than DNA for identification purposes."[1]

The career of the late evolutionary biologist Stephen Jay Gould was founded upon snails. Bahamian land snails of the genus *Cerion* were the subject of his actual field research, a focus he developed early in his career.[2] As his career advanced and began to be more concentrated on matters of theory building, he continued to return to the Bahamas regularly just to maintain his connection with the actualities and quotidian rewards of data collection that lie at the foundational core of science.

I guess you might say that, when it comes to the notion of expertise and its intersection with the law, handwriting identification cases have been my *Cerion* snails.

One day in early 1977, my longtime colleague, co-author, and friend Mark Denbeaux[3] came into my office and shoved a copy of *New York Magazine* at me. The

---

* The title is, of course, a portmanteau reference to three of my favorite titles: *Goodbye to All That*, a WWI memoir by Robert Graves and later an influential memoir essay by Joan Didion; *A Fool's Errand, By One of the Fools*, a novel by Albion W. Tourgee published in 1882, based on his experiences as an appointed judge in Reconstruction South Carolina attempting to help freedmen against the tide of the rising Ku Klux Klan; and *Dr. Strangelove, or, How I Stopped Worrying and Learned to Love the Bomb*, an influential 1960s movie directed by Stanley Kubrick. *See* Joan Didion, *Goodbye to All That*, in *Slouching Toward Bethlehem* (Dell 1967); Robert Graves, *Goodbye to All That* (Anchor Bks. 1927); Albion W. Tourgee, *A Fool's Errand, by One of the Fools* (Fords, Howard, & Hulbert 1878); *Dr. Strangelove, or, How I Stopped Worrying and Learned to Love the Bomb* (Sony 1964) (motion picture).

** John J. Gibbons Professor of Law, Seton Hall University School of Law. I would like to thank Charles A. Sullivan and Mark P. Denbeaux for helpful comments on a draft of this article, and Lesley Chenoweth Risinger for the usual thousand points of light, substantive and editorial.

1. Detective Steve White, testifying as a handwriting expert at trial in *Commonwealth v. Florence*, 120 S.W.3d 699, 701 (Ky. 2003).

2. *See* Stephen Jay Gould, *The Flamingo's Smile* 167–84 (W.W. Norton & Co. 1985).

3. Well, it's a long time now. Then it was only three and a half years.

magazine was open to an article entitled *Bruno Hauptmann Was Innocent*,[4] by the investigative journalist Anthony Scaduto. The article was largely based on a book by Scaduto[5] that re-examined the evidence against Bruno Richard Hauptmann, who was convicted in New Jersey's biggest criminal case ever, springing from the kidnapping and murder of Charles A. Lindbergh, Jr. in 1935. Scaduto had come to the conclusion that there were serious doubts about the fairness of Haupmann's trial, and the reliability of much of the evidence introduced in it. Part of his skepticism centered on the handwriting identification testimony of the prosecution's expert witnesses, most particularly that of Albert S. Osborn, the founding father of modern American handwriting identification expertise. Denbeaux asked me what I knew about handwriting identification testimony.

I am afraid my response was pretty conventional. I said that I didn't know anything about it, really, but that I assumed there was a pretty good basis for thinking that it was generally reliable, or else the courts wouldn't be admitting it. My naiveté pains me in retrospect.

In his usual mode as a natural deconstructionist, Denbeaux wasn't about to accept that—certainly not once I had said it was based on nothing but authority and trust. His (in retrospect justified) skepticism sent me scurrying to poke around a bit.[6] Poking around was not quite so easy in those days before computerized databases and the Internet, but what I found did not inspire confidence. I found no references to any research concerning the accuracy of the process, and it all seemed very subjective, and potentially extremely vulnerable to the power of suggestion from the other evidence in the case. Now Denbeaux and I were both skeptical, although still only on a relatively uninformed basis. So far it was all very casual, just one of those many interesting issues that cross the threshold of curiosity but usually go not much further. Then, one day in 1978, handwriting identification surfaced again, and this time it was not just a theoretical issue.

Denbeaux and I were doing quite a bit of real litigation in those days, alone and together, he more than I, really. It seemed like the right thing for young academics interested in the litigation system to keep their hands dirty with real cases, although today publication pressures might not as easily allow such a choice. The bigger cases were mostly civil rights actions, but there were a variety of other things. Denbeaux had been contacted by a friend who asked him to help the son of an acquaintance, a young man of around sixteen who had been employed in the summer of 1978 in a New York City Parks and Recreation program, which provided summer jobs for city youth. This particular youth had been assigned to keep time records and hand out paychecks to the other young people employed in his particular location in one of New York City's parks. A problem arose. A number of people complained that their checks had never arrived, and later it was discovered that they had been cashed after having been endorsed (in the payee's name). The endorsements were pretty clearly forgeries. Suspicion had fallen upon Denbeaux's client. The investigators demanded that the client provide handwriting

4. Anthony Scaduto, *Bruno Hauptmann Was Innocent*, N.Y. Mag. 60 (Nov. 22, 1976).

5. Anthony Scaduto, *Scapegoat: The Lonesome Death of Bruno Richard Hauptmann* (Putnam 1976).

6. General interest in the Hauptmann case led us down some other interesting paths, which included an interview in Hauptmann's cell, but that is another story.

exemplars for purposes of comparison with the forged endorsements.

We had toyed with the problems of handwriting identification for some time by reference to the analogous problem of eyewitness identification, and to Denbeaux what was proposed for his client was as objectionable as the client being requested to appear to be shown to witnesses alone in a show-up. He informed the investigating authorities that he would produce his client to provide exemplars only if exemplars were also taken from another appropriate group of writers of similar age and educational background, and all the exemplars were presented as candidates to the prosecution handwriting expert blind-coded, without the expert knowing who the target of the investigation was, or anything else about the case. (Denbeaux volunteered to obtain the "filler" exemplars from among our students, or from appropriate groups of high school students.)

And here is the surprising part. The young lawyer who was assigned to the investigation of the forgery case thought that this sounded reasonable. Like me in my initial naiveté, this lawyer assumed that there was something objective to handwriting identification, and that what was proposed would not really slow the process down much, or affect the result. He tentatively agreed to pursue the exemplars on the proposed basis, but later had to withdraw the offer when he discovered that the handwriting expert refused to have anything to do with the case under those conditions.

For a variety of reasons, not the least of which was this refusal by the handwriting expert, the investigation of Denbeaux's client petered out without any action being taken against him. But we had already begun to see more concretely the potential problems of claimed handwriting identification expertise.

As luck would have it, a handwriting identification issue popped up in one of the cases I was involved in not too long after. My client was the driver of a truck that was parked in a motel parking lot in New Jersey when, as the result of a tip from an informant, the truck was searched pursuant to a warrant. My client had driven the truck up from Florida, and he was charged with knowingly transporting the half-million Quaaludes that were found in the truck trailer. His position was that he was just a truck driver making a run, and that he knew nothing about the cargo. Certain written notes were found in the truck, the content of which indicated that the author of the notes was witting in regard to the nature of the cargo. The prosecution moved the court for an order to compel the production of handwriting exemplars by my client, so that they could be compared to the notes. I, of course, adopted the Denbeaux position from the New York summer job case.

Now this was litigation at a slightly different level than the New York case, which involved a pre-charge investigation of a relatively small matter by an overworked young lawyer in a juvenile justice unit. This was a major felony drug charge, the issues were raised post-indictment, the opponent was a smart and seasoned senior assistant prosecutor, and there was already a judge involved. And to say the least, the weight of precedent was against us. But we were blessed with a particularly intelligent judge, who managed to induce unrealistic expectations in me about the potentials of judicial performance in these matters for decades to come.

The judge read my papers carefully, considered the arguments about the unfairness of a show-up procedure in such a subjective area of claimed expertise, and ruled. He

ruled that, as the result of binding New Jersey precedent, he did not have the discretion to require the blind line-up procedure that we had requested, but that he would strongly recommend the prosecution to consider undertaking it voluntarily, because the reliability issues were significant. He indicated that he might revisit the issue of admissibility after the process was completed, and that at a minimum the prosecution risked reversal on appeal of any conviction obtained largely on the basis of such evidence generated by such a procedure. The stage was set for a real judicial examination of this issue.

And then the curtain crashed down, as it so often does. Other information indicating my client's knowledge of the drug transaction began to accumulate, and it quickly became time to cut a deal. I'm not sure the prosecution even got around to taking the exemplars, and the handwriting issue was totally washed out by the later evidence—it didn't even help get a better plea.

At the end of that case, I was interested in handwriting identification issues, but I still wasn't very knowledgeable. It didn't take a very deep study of handwriting identification theory and practice to write persuasively about its vulnerability to the dangers of "show-up" presentation of the evidence. And it didn't even take much depth of knowledge about the psychological literature concerning the potentially distorting effects of such suggestion. It was enough for purposes of litigation, or at least the litigation undertaken up to then, to take the judicial opinions that dealt with such subjects in the context of eyewitness identification and transfer them to the circumstances of handwriting identification.[7] But the Mayflower Madam changed all that.

In October of 1984, Cachet, an elite escort service in Manhattan, was shut down by authorities, and its principal, Sydney Biddle Barrows, was charged with promoting prostitution. Sydney was a Mayflower descendant and a member of the socially prominent Biddle family of Philadelphia, facts which added great popular interest to her case. She had previously been represented in regard to various business matters by a lawyer friend of Denbeaux's with whom he also had a consulting relationship. One thing led to another, and Denbeaux became Sydney's lead trial counsel. I had what was a relatively minor role in her defense, but it made me finally learn something in depth about handwriting identification expertise.

When the police shut Cachet down, they had seized a variety of handwritten records detailing customer identities, preferences, dates and times of service, charges, and so forth. The prosecution wanted to establish that those records were in Sydney's handwriting, and to that end, they moved to compel her to give handwriting exemplars. The case was prominent enough, and there were resources enough, to mount a full attack resisting the giving of such exemplars for any reason, or in the alternative, to require that they be given only in a blind process along with exemplars from other people of similar age, sex, and class, etc. I was in charge of this aspect of the case.

If this was going to be done right, we had to plan to explore the phenomenon of claimed handwriting identification expertise from the ground up: its history, its theory, its practice, and the empirical evidence that lay behind its claims to accuracy—if any.

7. Most particularly, *Wade* and its various progeny, federal and state. *See U.S. v. Wade*, 388 U.S. 218 (1967).

That entailed a serious research effort, part of which I did not feel completely competent to do on my own. That part was the part dealing with the empirical record—the discovery and evaluation of the formal research bearing on the reliability and validity of claims of handwriting identification expertise. For that I needed to bring in someone with credentials in research methodology as it bears on human skills claims. The first person I thought of was David Lykken, a University of Wisconsin Professor of Psychology who had recently (well, it was recent then) published a book on the empirical record concerning the claims surrounding the polygraph.[8] I called Professor Lykken. He couldn't do it, for a variety of reasons, but he gave me the name of someone he thought might be able to—a young associate professor at Boston College who had been doing really good research on jury behavior, and thus might be interested in issues with a legal connection. His name was Michael Saks.

Sometime in late 1984 or early 1985 I called him. He was interested. I took a train up to Boston and we had dinner. By the end of dinner we had a plan about how to frame the main questions and how to attack the literature search.

The Barrows case research was substantial and enlightening, but once again, it played only a small part in the case outcome. When exemplars were demanded from Ms. Barrows, we opposed on the same basis as in the previous cases, but this time we were able to document the paucity of research on the accuracy of handwriting identification. As in the New Jersey case, the judge ruled that he had no discretion under applicable precedent to require such a procedure, but he also indicated that he would allow the expert to be subject to properly designed blind tests of his claimed skills during cross examination. Subsequently, and perhaps as a result of the latter ruling, the people's expert rendered a report that indicated that his (non-blind) attribution of authorship of various documents to Ms. Barrows was only "probable," an opinion that the prosecutor found less than helpful.

After a large number of twists and turns to which I was mostly a spectator, Sydney was finally allowed to plead to a misdemeanor and pay a fine, a plea deal that Denbeaux later characterized as a "kiss on the wrist." But all of our research did not go to waste. After I returned from a visitorship at the National University of Singapore and finally re-adjusted to life in New Jersey, we spent the better part of a year finishing the research and drafting the text that finally became the first article to take an in-depth look at the tenability of the claims of any area of forensic identification "science": *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification "Expertise."*[9]

At this remove, that article is probably most remembered (by whoever remembers it at all) as an article about handwriting expertise, and so it was. Having surveyed and combed the extant literature for any trace of formal empirical evidence for the accuracy of the skills claimed by handwriting identification experts, we found that there was either none, or that (depending on how you counted the results of certain proficiency tests), it

8. David Thoreson Lykken, *A Tremor in the Blood* (McGraw-Hill Bk. Co. 1983).

9. D. Michael Risinger, Mark P. Denbeaux & Michael J. Saks, *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification "Expertise"*, 137 U. Pa. L. Rev. 731 (1989).

was unsupportive of the claims of the field.

And so the article called into question the rationality of allowing such unvalidated testimony from handwriting identification experts. But we hoped that it would have a much wider effect. By this point, we had come to suspect that the accuracy claims of a large number of "forensic science" enterprises were epistemically weak. We hoped by enlarging on the themes suggested by the notion of "exorcism" invoked in the title, to give the article broader significance. We did our best, but I suspect for most people it has remained "that handwriting article." However, we viewed it as the beginning of a process for the re-examination of the very notion of expertise, and the criteria by which expert validity is to be judged, especially when it is proffered by the prosecution in a criminal case. In a way, we regarded handwriting identification as one of the weaker forms of such asserted expertise, and one of the easiest to use to illustrate the kinds of weakness applicable to other members of the genus. The larger themes of the article were put thus in the conclusion:

> We have reached the end of our main topic and it only remains for us to ask if there are any greater lessons to be learned from the law's treatment of handwriting expertise. One is tempted to say that as a general proposition, the law does not yet know how to deal with science, or with things asserting themselves to be science. On the one hand there is the *Frye* test and its variants, which seem to incorporate a requirement of validity testing from the scientific community as a condition of admissibility of evidence claiming the mantle of science. On the other hand, the *Frye* test has been criticized as denying the law the benefits of new developments and has been grudgingly construed to apply only to forms of expertise that are both centrally scientific and novel. Thus, alleged areas of expertise, whose practitioners eschew the label "scientific", are sometimes accepted without evidence of validity; so are forms of expertise that predated the *Frye* test. A lot of unvalidated nonsense is allowed into court under circumstances that look suspiciously close to the old gentlemen's agreement from the days of Langdellian-Wigmorean "juridical science": "I won't look too closely at your claims to 'scientific' expertise if you don't look too closely at mine."
>
> Finally, beyond the dark corners of inherited expertise, we may perhaps draw a broader lesson still. The "search for truth" model of our procedural system, which comprises much of its official ideology, cannot account for many phenomena in the real world of litigation. Many have therefore proposed an alternate lens for viewing the reality of litigation, the "judicial combat" model. The proponents of this view assert that what we actually do (as opposed to what we say we do) is seek, not a result that mirrors the law applied to the facts as they were, but a result that is the product of a fair and satisfying contest, whatever the facts may have been. If this latter view is not sufficiently cynical as an account of many of the details of actual practice, we are afraid that the handwriting expertise phenomenon, and others, lead us to propose still another model: the "exorcism of ignorance" model. This model holds that, under the stress of having to wrestle with important types of facts about which there is no good evidence, the system will invite the creation of a proxy for rational knowledge, a form with the appearance of evidence but no rational content, to be used in a ritual exorcism of an ignorance we cannot bear. If this in fact accounts for some of what happens in the real world, we can only hope that the legal system's conscious commitment

> to the "search for truth" model will lead it to respond honorably when the imperfection in our knowledge is finally pointed out in ways that cannot be avoided.[10]

In many ways, the years since have been devoted to trying to provide the required information, and to holding on to hope.

Early on, I guess that, at least in regard to handwriting identification, we expected that maybe our article would precipitate some action in the courts, or some published reaction from the handwriting community. We were not quite naïve enough to think that judges and criminal defense attorneys read law review articles for amusement, but we figured that maybe when a defense attorney had a case where the testimony of a handwriting expert was going to convict his client, he might do enough research to find an extensive analysis of the subject in a sixty-page article in the August University of Pennsylvania Law Review. After all, Lexis and Westlaw were by this time in operation (although I suppose most criminal defense attorneys did not have access to them). We did get some gratifying response from the community of evidence scholars,[11] but from the handwriting community, the courts, and the criminal defense bar, for years nothing.[12] Then came *Daubert*.[13]

It is commonly said that *Daubert* worked a revolution. In a sense it did, but in another more important sense, it is still too early to tell, or at least too early to tell what the endpoint will be, and the extent to which it will be revolutionary. In fact, it is likely to take so much time to work out that the notion of "revolution" will be an uncomfortable fit for the process or the result. That is probably so already. And in some areas, it could end up having little practical effect at all.

Don't misunderstand me. *Daubert* is definitely an important case, if for no other reason than that it catalyzed a discussion of the nature of expertise and what constitutes reliable expertise which had just begun in the academic (and polemical) communities,[14] but which had not yet penetrated the cogitations of the courts, with a couple of notable and influential exceptions.[15] And *Daubert* itself laid the foundation for the much more

---

10. *Id.* at 779–82 (footnotes omitted).

11. *See e.g.* Kenneth R. Kreiling, *Scientific Evidence: Toward Providing the Lay Trier with Comprehensible and Reliable Evidence Necessary to Meet the Goals of the Rules of Evidence*, 32 Ariz. L. Rev. 915 (1990); Roger C. Park, *Evidence Scholarship, Old and New*, 75 Minn. L. Rev. 849 (1991).

12. Professor Giannelli and his co-author Carin Cozza describe the article as "remaining below the judicial radar" until the *Daubert* decision. *See* Paul C. Giannelli & Carin Cozza, *Forensic Science:* Daubert *Challenges to Handwriting Comparisons*, 42 Crim. L. Bull. 347, text at n. 14 (2006).

13. *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993).

14. Many academic commentators concentrated on the weaknesses of expertise in the criminal case setting. *See e.g.* David L. Faigman, *The Battered Woman Syndrome and Self Defense: A Legal and Empirical Dissent*, 72 Va. L. Rev. 619 (1986); Randolph N. Jonakait, *Real Science and Forensic Science*, 1 Shepard's Expert & Sci. Evid. Q. 435 (1991); Risinger et al., *supra* n. 9. Courts often seemed most troubled in civil cases, perhaps beginning with Judge Weinstein's decision in *In re "Agent Orange" Prod. Liab. Litig.*, 611 F. Supp. 1223, 1253 (E.D.N.Y. 1985) (characterizing plaintiff's expert proffers on Agent Orange's cancer causation as "conclusory and subjective"). *See generally* Michael D. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of* Agent Orange *and Bendectin Litigation*, 86 Nw. U. L. Rev. 643 (1992). The polemicist, of course, is Peter Huber, whose 1991 book *Galileo's Revenge: Junk Science in the Courtroom* may have had an impact in framing issues, stirring public consciousness, and motivating judicial action all out of proportion to the balance (or lack thereof) of its presentation.

15. *See e.g. In re "Agent Orange,"* 611 F. Supp. 1223. See also *U.S. v. Downing*, 753 F.2d 1224 (3d Cir. 1985), an influential opinion by Judge Becker (cited in *Daubert*) that reversed the rejection of defense proffered expertise on eyewitness identification weaknesses, and laid out an influential reliability framework,

sophisticated *Kumho Tire* opinion. But *Daubert* is in many ways Delphic,[16] and in others deficient. Given this (which I will expand on a bit below), it is surprising, to me at least, that people still commonly use the term "*Daubert*" as a label suggesting a coherent or at least determinant set of doctrines.

However, whatever its conceptual strengths or weaknesses, the *Daubert* decision electrified litigators. By virtue of its very weaknesses, the opinion threw open the possibility of challenge wherever expertise was proffered (at least in a federal court), and that was everywhere. *Daubert* motions became the new flavor of the month, especially in federal civil litigation. But the possibilities were not completely lost on the federal criminal defense bar, either. However, in the first year and a half after the decision in *Daubert*, things were slow to get started, and it was not clear what the focus of criminal defense invocation of *Daubert* was going to be. There appeared to be much more effort expended trying to use it to gain admission for defense proffers, particularly in regard to polygraph results, than to challenge the reliability of prosecution proffers.[17] Finally, in 1995 a full-scale challenge to a prosecution proffer of "forensic science" testimony was raised, and when it did finally come, the target was handwriting identification.

This was hardly surprising. Handwriting identification was perhaps no more unreliable than many forensic identification techniques, such as visual hair comparison or bite mark comparison, which have been shown to have miscarried in many of the DNA exonerations established in the past decade.[18] But what handwriting identification had that the other forensic identification disciplines lacked was a full literature search documenting the unverified nature of the expert claims involved. So the presence of this

---

id. at 1233–37. The fact that Judge Becker's opinion was pro-admissibility was one of the reasons that some early commentators predicted that *Daubert* was intended to, and would, lead to more relaxed standards of expert admissibility.

16. It was not even completely clear from the *Daubert* opinion itself whether it was intended to tighten or relax standards of expert admissibility. This was reflected both in popular sources and in academic commentary. For popular sources, consider these two articles, published the day after the *Daubert* decision: Compare Joan Biscupic, *Judges Get Broader Discretion in Allowing Scientific Testimony*, 116 Wash. Post A6 (June 29, 1993) (stating that the Supreme Court relaxed standards for admission) with Linda Greenhouse, *Justices Put Judges in Charge of Deciding Reliability in Scientific Testimony*, 142 N.Y. Times A13 (June 29, 1993) ("The 7–2 decision invited judges to be aggressive in screening out ill founded or speculative scientific theories."). For academic commentary, 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 raised the 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).

17. Most of the early opinions generated in criminal cases involving *Daubert* issues involved criminal defendants attempting to obtain the admission of polygraph results. *See e.g. U.S. v. Black*, 831 F. Supp. 120 (E.D.N.Y. 1993); *U.S. v. Rodriguez*, 37 M.J. 448 (Armed Forces App. 1993); *see generally* D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?* 64 Alb. L. Rev. 99, 128–31 (2000) (analyzing the boom and bust in polygraph cases).

18. *See* Brandon L. Garrett & Peter Neufeld, *Report to the Committee on Identifying the Needs of Forensic Science [National Academies of Science/National Research Council]: The Incidence of Improper Forensic Science Testimony in Criminal Trials of the Innocent* (2008) (unpublished, copy on file with the author). This study examined the trial records in 114 of the 138 trials (in the first 212 cases that later resulted in DNA exonerations) that were supported by "forensic science" evidence, and found that microscopic (visual) hair comparison and bite mark analysis (especially hair comparison) were among the leading sources of such evidence contributing to such miscarriages. *Id.* at 2 (abstract). Incidentally, the absence of handwriting identification cases in the DNA exonerations is likely more a reflection of the kind of case in which DNA is available, rather than an indication of the high quality of handwriting identification evidence.

documentation in *Exorcism* made handwriting identification the easiest *Daubert* target.

The bull's eye on the back of handwriting identification was not lost on the boffins at the FBI laboratory, where it was apparently thought prudent to commission some research directed toward giving some sort of empirical warrant to the existence of a general handwriting identification skill inhering in government laboratory document examiners, and in 1994 the first fruits of that effort were published in the Journal of Forensic Sciences: *Proficiency of Professional Document Examiners*, by Moshe Kam, Joseph Wettstein, and Robert Conn.[19] The methodological oddities and limitations of this and subsequent handwriting identification research by Professor Kam and his associates have been the subject of fairly intense controversy ever since,[20] but I must admit that I did not know much about this study until I became involved in the handwriting challenge in the *Starzecpyzel* case in late 1994.

I had nothing to do with the underlying case in *United States v. Starzecpyzel*,[21] nor with the decision to mount a *Daubert* reliability challenge. The defendants were represented by private counsel, Alan Haber and Michael Hurwitz, who approached Barry Scheck of Cardozo Law School to see if Cardozo clinics might be interested in helping in a full-scale *Daubert* challenge to the handwriting identification that was at the heart of the government's case against the defendants. Ultimately, Mira Gur-Arie, now Director of the Interjudicial Relations Office at the Federal Judicial Center, but then a clinician at Cardozo, took over as lead attorney in mounting the challenge, and she contacted us.

By us, I mean Michael Saks and myself. Michael was retained to testify in the *Daubert* hearing about the processes of science and the results of our search for validity data regarding claimed handwriting expertise. I was a consultant, which meant mostly I got to sit and watch while Mira cross-examined the main government witness, Mary Wenderoth Kelly, then chair of the Committee on Testing of the American Board of Forensic Document Examiners.[22] By the end, I did not see how we could lose.[23]

We lost.

Well, only sort of, in retrospect. There had been people warning me that, since it was not clear what the requirements of *Daubert* were for expertise that was not the product of science, we might lose on that basis. But I just could not bring myself to believe that an area of claimed expertise that had called itself a "science" for nearly a

19. Moshe Kam et al., *Proficiency of Professional Document Examiners in Writer Identification*, 39 J. Forensic Sci. 5 (1994).

20. *See* Appendix, *infra* nn. 53–88 and accompanying text. For a fuller discussion see D. Michael Risinger, *Handwriting Identification*, in Faigman et al., *Modern Scientific Evidence* §§ 33:28–33:31 (West 2008).

21. 880 F. Supp. 1027 (S.D.N.Y. 1995).

22. The American Board of Forensic Document Examiners (ABDFE) is the most prestigious credentialing body in the field of document examination. Whether that means the people they certify are more skilled or accurate than other document examiners is a different question. For a thumbnail sketch of the training and credentialing problem, see Appendix, *infra* notes 7–10 and authorities there cited. For a brief discussion of the problems in confusing credentials with a reliability warrant (and the problem of astrology), see Appendix, *infra* note 115.

23. Ms. Kelly's testimony is summarized in *Starzecpyzel*, 880 F. Supp. at 1031–35. One thing of interest that was not included in the court's summary was that her testimony established just how weak the certifying criteria of the American Board of Forensic Document Examiners were at that time. See the discussion in D. Michael Risinger & Michael J. Saks, *Science and Nonscience in the Courts:* Daubert *Meets Handwriting Identification Expertise*, 82 Iowa L. Rev. 21, 38 nn. 94, 96 (1996), and sources there cited. It is not clear if

century, and that operated out of places like the FBI "forensic science" laboratory, would be allowed to escape validity testing on that basis.[24] And in his opinion Judge McKenna eviscerated any remaining pretensions handwriting identification witnesses might have had to scientific status.[25] But that was, ironically, their ticket into court in *Starzecpyzel*. Because the products of this "forensic science" were not based on science, they were unaffected by the requirements of *Daubert* (which they concededly could not meet) and thus were admissible as they had always been.[26]

If I was a bit disheartened by Judge McKenna's opinion, the document examiner community was not elated by it either. It was hardly a ringing endorsement, and it was certain that more challenges would be coming—indeed, two were already in the pipeline. While the decisions in those cases[27] did nothing that should have threatened the document examiner community too much, it seemed that the time was ripe for more exciting things to come.

At about this time, David Faigman organized an effort that largely determined my fate as an observer of the handwriting identification species of snails. He proposed to West Publishing Company a new treatise dealing with the scientific strengths and weaknesses of various common forms of expert evidence in litigation, to be edited by himself and David Kaye, Michael Saks, and Joseph Sanders, and to be called *Modern Scientific Evidence*. These editors asked me if I would agree to write the chapter on handwriting identification. Given the fact that litigation over handwriting identification reliability seemed to be heating up, and that it seemed to be the subject most likely to be involved in trail-breaking for the issues involved in prosecution-proffered expertise, this seemed like a great opportunity, so I agreed to do it. And so I have, from mid-1995[28]

---

24. I had not yet become acquainted with how desperately the bulk of trial judges do not want to exclude "forensic science" testimony that the prosecution wants in, virtually no matter what. See the discussion *infra* notes 111–18 and accompanying text.

25. *Starzecpyzel*, 880 F. Supp. at 1036–38. He concludes, "In sum, the testimony at the *Daubert* hearing firmly established that forensic document examination, despite the existence of a certification program, professional journals and other trappings of science, cannot, after *Daubert*, be regarded as 'scientific . . . knowledge.'" *Id.* at 1038.

26. *Id.* at 1042–47. This summary is not really fair to the subtlety with which Judge McKenna dealt with the reliability decision under 702 once outside of the *Daubert* framework. He explicitly rejected the relevance of the "long history of judicial acceptance." *Id.* at 1044 n. 20. His "two stage" analysis of the process of handwriting examination laid the foundation for the later decisions in *McVeigh* (Appendix case 5) and *Hines* (Appendix case 8), and he even considered adopting the *Hines/McVeigh* result himself. *Id.* at 1047. In addition, he even anticipated *Kumho Tire v. Carmichael* in his limitation of his decision only to the task in front of him, which was determining the genuineness of challenged signatures when there was a large amount of signature material to work with. *Starzecpyzel*, 880 F. Supp. at 1043. And his suspicion of the appearance of false precision in the document examiners "nine point scale" of conclusions was salutary, as was his willingness to restrict the form of testimony that could be given and to fashion a cautionary instruction to combat over-reliance by the jury. I have some criticisms, of course, mainly that Judge McKenna got the burden of persuasion on the issue of reliability wrong. *See* Risinger & Saks, *supra* n. 23, at 31–32 and especially n. 69. (One of the effects of the Supreme Court's opinion in *Kumho Tire v. Carmichael* was to clarify that, after the opponent of a proffer of expertise has shown that there is a tenable claim that the proffer is unreliable, the burden is on the proponent to establish sufficient reliability to satisfy the requirements of Federal Rule of Evidence 702. *See Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137, 149, 152 (1999). *See also* Appendix, *infra* n. 105.) Nevertheless, considering what I have seen since, if every subsequent court had performed as well as Judge McKenna, we would not be where we are.

27. *U.S. v. Velasquez*, 64 F.3d 844 (3d Cir. 1995) (Appendix case 3); *U.S. v. Ruth*, 42 M.J. 730 (Army Crim. App. 1995) (Appendix case 2).

28. The first edition was not actually put to bed until late 1996 and bears a 1997 copyright date.

until the present (March, 2008). I was hoping to see the American judiciary, most particularly the federal judiciary, perform that magic of working out the details of reasonable doctrines that is sometimes known as "the genius of the common law."[29] As it turns out, I saw something different. One must not forget that "[t]he life of the law has not been logic: it has been experience."[30]

The years from mid-1995 until 1999 were a little disappointing. It did not appear to me that most of the seven courts that considered handwriting reliability issues (six federal, two from the trial level and four appellate,[31] and one state appeals case[32]) were taking them seriously enough to actually frame and determine the real issues presented by the cases. In addition, in late 1997, the Supreme Court decided *General Electric v. Joiner*,[33] which signaled potential trouble ahead in expert reliability jurisprudence generally. How much it would affect my little Cerions I would not have predicted. Let us delay consideration of the effects of *Joiner*, however, while we look generally at the cases from this period.

THE CONTOURS OF THE CASES, 1995–1999

As I just said, the seven handwriting reliability cases from mid 1995–1999 were somewhat troubling. Indeed, I consider the Sixth Circuit opinion in *United States v. Jones* to be one of the worst crafted and reasoned opinions I have ever encountered in my forty plus years of reading opinions.[34] I have already discussed *Starzecpyzel* above, and both *Starzecpyzel* and the others (including *Jones*) are fully analyzed and discussed in the Appendix, so I will limit myself to general observations here. There were six federal cases comprising two district court decisions (*Starzecpyzel* and *McVeigh*) and four by appellate courts (*Ruth*, *Velasquez*, *Gonzales*, and *Jones*). The two district court decisions were given after extensive hearings, and both Judge McKenna and Judge Matsch were clearly troubled by the weakness of the warrant for handwriting identification reliability in general. Judge McKenna limited his decision to admit the testimony (perhaps luckily, perhaps with intuitive prescience) to the single issue in front of him, signature authentication, and also fashioned a cautionary instruction designed to take the aura of "science" away from the testimony.[35] Judge Matsch, in a determination that was destined to become quite influential even though it was never formally published, restricted the testimony of the document examiner to pointing out similarities and differences between the known and questioned writing, but forbade testimony giving

29. *See* Sir Frederick Pollock, *The Genius of the Common Law* (Columbia U. Press 1912).

30. Oliver Wendell Holmes, *The Common Law* 5 (Mark DeWolfe Howe ed., Harvard U. Press 1963). This phrase has always been looked upon as both pregnant and enigmatic. *See e.g.* Susan Haack, *On Logic and the Law: "Something but Not All"*, 20 Ratio Juris. 1 (2007); Max Radin, *Law as Logic and Experience*, 54 Harv. L. Rev. 711 (1941).

31. *See* Appendix: case 1, *Starzecpyzel*, 80 F. Supp. 1027; case 2, *Ruth*, 42 M.J. 730; case 3, *Velasquez*, 64 F.3d 844; case 4, *U.S. v. Gonzales*, 90 F.3d 1363 (8th Cir. 1996); case 5, Pre-Trial Transcr., *U.S. v. McVeigh*, 1997 WL 47724 (D. Colo. Feb. 5, 1997); case 6, *U.S. v. Jones*, 107 F.3d 1147 (6th Cir.1997).

32. Appendix: case 51, *State v. Cochran*, 1997 WL 399835 (Ohio App. 4th Dist. July 8, 1997).

33. 522 U.S. 136 (1997).

34. My analysis is fully set out in the Appendix case 6.

35. See the discussion infra at 26.

a conclusion about common authorship.[36]

However, the four appellate decisions seemed to come from an unrelated universe. Except for *Jones*, they did not involve extensive hearings at the trial level, and neither the reviewing court nor the trial court (from what can be gleaned in the opinions) seemed to care about hard issues about evaluating evidence concerning reliability. Instead, all four simply accepted the self-believing assertions of the proffered experts as sufficient to establish their reliability. In other words, they applied what I have elsewhere called "the guild test."[37] And the single state court decision (from the Ohio Court of Appeals) also fit this mold. This mold, I feared, would be the shape of things to come.

THE TROUBLE WITH *DAUBERT*

I have already indicated reservations about *Daubert* as an opinion. I am hardly alone.[38] On the one hand I think that on a very general level, *Daubert* was both necessary and beneficial. One does not have to be a fan of Peter Huber[39] to conclude that by the early 1990s in many areas of American litigation practice the accuracy-distorting effects of commonly used faux expertise required some systemic response. And if that response was going to be provided through court decision, it would have to either be facially trans-substantive, or else whatever contextual nuances it might properly manifest were going to have to be spelled out and explained. And this was going to be a seriously difficult task for courts, both because of the well known difficulties of non-experts judging expert claims,[40] and because some of the expert claims involved had, properly or improperly, assumed the honorific title of "scientific" claims. Into this mess, Justice Blackmun and the *Daubert* majority boldly strode.

Various weaknesses of the resulting opinion have been pointed out by others.[41] For me, and the epistemic health of my little Cerions, the weaknesses were basically threefold. First, it was not clear from the opinion whether the general intendment was to allow more in or to keep more out. One could (and later some did) put forth a

---

36. The oral argument on the issue is reported at *McVeigh*, 1997 WL 47724, and Judge Matsch's oral decision is *id.* at *23. The colloquy, with Judge Matsch's questions and observations, is more sophisticated than most written opinions in this area have been.

37. The contours and implications of the "guild test" and the entailed "sufficient experience" test were worked out in D. Michael Risinger, *Defining the "Task at Hand": Non-Science Forensic Science after* Kumho Tire v. Carmichael, 57 Wash. & Lee L. Rev. 767, 770–78 (2000).

38. For general criticisms of the coherence of *Daubert*'s approach, see generally David S. Caudill and Richard Redding, *Junk Philosophy of Science?: The Paradox of Expertise and Interdisciplinarity in Federal Courts*, 57 Wash. & Lee L. Rev. 685, 691–92 (2000) and authorities collected therein at note 17.

39. *See supra* n. 14.

40. *See generally* Ronald J. Allen & Joseph S. Miller, *The Common Law Theory of Experts: Deference or Education?* 87 Nw. U. L. Rev. 1131 (1993); Scott Brewer, *Scientific Expert Testimony and Intellectual Due Process*, 107 Yale L.J. 1535 (1998). Professor Brewer concludes that ordinarily non-scientist judges and juries can never be warranted in evaluating the claims of disagreeing scientists and rationally choosing between them based on such a non-expert evaluation. This seems a little extreme as a generalization. A later piece, *Navigating Uncertainty: Gatekeeping in the Absence of Hard Science*, 113 Harv. L. Rev. 1467 (2000), persuasively argues that there are choice strategies in some cases of "strong [scientific] uncertainty" which are not merely decisions about deference. *Id.* at 1470. And of course, the more an area of expertise approaches common experience, the more ways of rationally evaluating plausibility a factfinder may have. *See* Roger Park, *Signature Identification in Light of Science and Experience*, 59 Hastings L.J. ___ (forthcoming 2008).

41. *See supra* n. 98.

sophisticated rationale which resulted in a little of both,[42] but it is stretching things to say that Justice Blackmun foresaw this outcome. While it seemed to me that the actual result of *Daubert* dictated that it was dominantly intended to raise the admissibility bar (and that seems now to be clear from later decisions of the Supreme Court itself[43]), the existence of the ambiguity allowed courts to resist exclusionary results as not comporting with *Daubert*'s intendment.[44]

Second, as noted earlier, it was not clear whether the general emphasis in *Daubert* on judicial responsibility to pay serious attention to claims of expert reliability or lack thereof under Rule 702 was intended to apply to all expertise, or to merely "scientific" expertise, whatever the demarcation between the two might be. (The assumption that *Daubert* was just for real science was one of the basic rationales of the decisions in both *Starzecpyzel* and *Ruth*, and, somewhat more ambiguously, in *Velasquez* and in *Jones*.)

Third, there were those damned *Daubert* factors. They are often summarized as four factors,[45] and set out something like this from the *Jones* case:

---

42. *See* Faigman et al., *supra* n. 16, at 1810 (arguing that the Federal Rules would let in strongly grounded novel claims that *Frye* would not, but would keep out weakly grounded claims that had gained acceptance among some reference class of practitioners that might have been taken to satisfy *Frye*). This is expanded upon in Faigman et al., *supra* n. 20, at § 1:14.

43. This result seems to have been accepted as a matter of course in both *General Electric v. Joiner* and *Kumho Tire*, but it is most clearly embraced in the following passage from Justice Ginsburg's opinion for a unanimous court in *Weisgram v. Marley*, 528 U.S. 440, 442 (2000) (speaking on the issue of a Court of Appeals remanding for entry of judgment instead of new trial when proffered expertise is found wanting on appeal). "Since *Daubert*, moreover, parties relying on expert evidence have had notice of the exacting standards of reliability such evidence must meet." *Id.* (I have heard Paul Giannelli refer to *Marley* as the fourth case in the *Daubert* Trilogy.).

44. Startlingly, one of the handwriting reliability opinions from 2003 relies upon this characterization of *Daubert*. *See U.S. v. Jabali*, 2003 WL 22170595 at *2 (E.D.N.Y. Sept. 12, 2003) (Appendix case 35).

45. As Professor Denbeaux and I have said:

> Referring to "four" factors has become standard, though the real number of factors is subject to debate. The *Daubert* opinion spake thus, without numbering factors: "a key question [in regard to a theory or technique] . . . will be whether it can be (and has been) tested." 509 U.S. at 593. "Another pertinent consideration is whether the theory or technique has been subjected to peer review and publication. Publication (which is but one element of peer review) is not a *sine qua non* of admissibility . . . ." *Id.* "Additionally, in the case of a particular scientific technique, the court should consider the known or potential rate of error . . . and the existence and maintenance of standards controlling the technique's operation." *Id.* at 594 (internal citations omitted). "Finally, 'general acceptance' can yet have a bearing on the inquiry." *Id.* These were summarized in *Kumho Tire* as "several factors" without numbering, but with four bullet points. 526 U.S. at 149–50. However, it is easy to separate whether a claim "can be tested" (its empirical nature or theoretical falsifiability) and the degree to which it has been subjected to actual testing, into two separable but nested factors. In addition, the potential rate of error is arguably always 100 percent in the absence of some kind of testing (though not necessarily the kind of formal testing that would lead to more specific and quantifiable knowledge of an error rate). Knowledge of error rates is thus a product of testing. In addition, can "standards of control" for a technique's operation be a relevant factor if there is no reason to believe such "standards" enhance reliability? This too would seem to be a question of testing, at least in some contexts. Finally, a fortiori "general acceptance" is the product of peer review, so one can argue that there are really eight explicitly referenced "*Daubert* factors" (falsifiability, testing, peer review, publication, potential error rate, known error rate, standards of practice, general acceptance) or only three (falsifiability, testing which reveals error rate, peer review). In addition, the *Daubert* Court invokes the relevance-based concept of fit, 509 U.S. at 591, which is perhaps best seen as an analogue to "external validity," and which can easily be asserted as a fifth (or ninth, or fourth) "*Daubert* factor." Courts have not always referred to four *Daubert* factors, either. *See e.g. U.S. v. Crisp*, 324 F.3d 261, 266–67 (4th Cir. 2003) (five factors); *U.S. v. Prime*, 220 F. Supp. 2d 1203, 1204 (W.D. Wash. 2002) (five factors); *U.S. v. Griffin*, 50 M.J. 278, 284 (Armed Forces Ct. App. 1999) (six factors).

(1) "whether a theory or technique . . . can be (and has been) tested"; (2) "whether the theory or technique has been subjected to peer review and publication"; (3) "the known or potential rate of error"; and (4) "general acceptance."[46]

First, one should note that what the Court actually said, which is set out in footnote 45, is much more extensive and nuanced than the summary. In addition, the rest of the *Daubert* opinion expands quite a bit on what is intended by various factors, especially in regard to the role of publication in the general process of peer review of which its publication is only a part, and points out that publication does not insure reliability.[47] However, in the hands of many courts, something like the *Jones* version of the "*Daubert* factors" has become a mechanical checklist, woodenly applied. Perhaps it is not Justice Blackmun's fault that many lower courts have tried to make a code of regulations out of a caricature version of the "*Daubert* factors." But the utility of the opinion has been undermined by this tendency among the lower courts. It has even surfaced in handwriting identification reliability cases,[48] though not in the current pre-1999 group, when the emphasis was dominantly on declaring the factors largely irrelevant because handwriting identification is not "scientific" expertise.

The fact that the *Daubert* factors have been applied in a ridiculously mechanical way by many lower courts does not diminish the fact that what the Court set out has problems of its own, perhaps most pertinent for our purposes in regard to the notion of "error rates," about which factor the Court says least.[49]

One problem with the "error rates" formulation is the contrast between "known" and "potential" error rate. If the subject matter of expert testimony suggests some conclusion to be drawn from the testimony by the jury (and it would be hard to imagine testimony that did not so suggest in some way that would still count as relevant), then one way to look at error rates would be to ask how often the suggested conclusion would be right (correspond with the facts of the world) and how often wrong. For any

---

Mark P. Denbeaux & D. Michael Risinger, Kumho Tire *and Expert Reliability, How the Question You Ask Gives the Answer You Get*, 34 Seton Hall L. Rev. 15, 32 n. 64 (2003) (internal jumpcites omitted). The Advisory Committee's note for the 2000 revision of Rule 702 lists five factors, but then adds five more that were derived from intervening case authority. The Advisory Committee note is a pretty good piece of work, except for a couple of details. For me, the most irritating is that they selected *U.S. v. Jones* as a case citation for the proposition that experience and training alone could form the basis of reliable expertise. While the proposition for which the case was cited is in its general form uncontroversial, it has been argued that the citation to *Jones* put the Advisory Committee's seal of approval on the reliability of handwriting identification expertise globally. *See U.S. v. Ojeikere*, 2005 WL 425492 at *3 (S.D.N.Y. Feb. 18, 2005) (Appendix case 41). This was unlikely to have been the intendment of the Committee, but the opinion in *Jones* is such a mess in its own right that I am tempted to believe it was selected by glancing at its headnotes. Better they should have cited *Velasquez*, which at least has the virtue of accurate and relatively logical exposition.

46. This particular version of the summary is taken directly from *Jones*, 107 F.3d at 1156.

47. 509 U.S. at 593.

48. This seems to have been less of a problem in handwriting opinions than others, perhaps because the usual approach to the factors fits the conditions of handwriting expertise so poorly. However, even handwriting opinions have their examples of forcing handwriting pegs into the *Daubert* checklist holes. *See e.g. U.S. v. Gricco*, 2002 WL 746037 (E.D. Pa. Apr. 26, 2002) (Appendix case 23) (applying the "*Paoli II* factors" derived from *In re Paoli R.R. Yard PCB Litig.*, 35 F.3d 717 (3d Cir. 1994), which include the *Daubert* factors). The court shows it knows the words, but not the tune. See also the appellate opinion in *U.S. v. Prime* (*Prime II*), 363 F.3d 1028 (9th Cir. 2004) (Appendix case 38).

49. Similar observations might be made about, for example, the mysteries of the relationship of testability and testing, but the illustration in regard to "error rates" is sufficient for present purposes. For more general criticisms of the coherence of *Daubert*'s approach, see generally Caudill & Redding, *supra* n. 38, at 691 n. 17.

testimony for which there is no reason to have any notion of the magnitude of the "wrong versus right" ratio, the "potential" error rate is a fortiori 100%. Once we get information that starts to suggest at least some limits on this, the error rate starts to become known, and the notion of "potential error rate" is defined by the information putting limits on the initial 100% potential. To make a long story short, the "potential error rate" is in this view the same as the "known error rate" whenever there is any knowledge warranting the conclusion that the potential error rate is not 100% (or, more precisely, random correspondence).

Or perhaps the notion of "known error rate" was reserved for circumstances where there is good empirical evidence that would suggest a fairly mathematically determinate error rate, and the notion of potential error rate was to cover limits of error rates known by reference to more qualitative and less quantitative sources of information, which establish broad ranges of potential error which are indeterminate at their specific boundaries.

Whatever the reader's response to these issues (and there are others), you can see that the casual invocation of these concepts by the Supreme Court was unlikely to provide much guidance to the lower court in how to use the notion of error as a part of a test for reliability.[50]

THE TROUBLE WITH *GENERAL ELECTRIC V. JOINER*[51]

I was reflecting on such issues, and analyzing the slow evolution of my study species (including the malformed *Jones* opinion) when, in December of 1998, the Supreme Court decided *General Electric v. Joiner*. From my perspective, the case had two main points: It made clear that Rule 702 reliability decisions were to be reviewed on appeal under an abuse of discretion standard,[52] and it relaxed the apparent prohibition on considering particular results of application in determining reliability.[53] The latter attempt to hermetically seal off reliability of process from reliability of product was one of *Daubert*'s more dubious and less workable conceptual distinctions, and the Court in *Joiner* recognized that often a proper determination of Rule 702 reliability required taking the whole process in question in the case into account. I found this heartening, since one of my conclusions from the study of both the handwriting cases, and others, was that proper reliability determinations had to be directed toward the subtask of an expertise that was actually being undertaken in the particular case,[54] and I believe *Joiner* made such an analysis potentially more attractive to courts. But I was worried about the other part of *Joiner*.[55] Under an abuse of discretion standard, the courts of appeals will

50. In my opinion, the most common error of courts in dealing with prosecution proffers of claimed forensic science expertise has been to approach the notion of "error rates" globally instead of task-specifically, even after *Kumho Tire*, and to accept low quality evidence on the incidence and magnitude of error in various processes.

51. 522 U.S. 136.

52. *Id.* at 142–43.

53. *Id.* at 146–47.

54. The necessity of task-specific testing of error rates in any process of evaluation of expert claims was a theme of Risinger et al., *supra* note 9. *See id.* at 742–43, 750.

55. And again, I was not alone. See, e.g., David L. Faigman, *Appellate Review of Scientific Evidence under*

generally end up affirming whatever the district court decided as long as it passed some very low threshold of tenability. That would give the court of appeals very little authority to norm the district court judges in their circuit into a uniform approach to reliability issues, and it would also give court of appeals opinions very little claim to the status of precedent, or even (given the casual context in which the decisions were necessarily generated) to being regarded as products of sufficient reflection to count as serious authority. At least as far as the handwriting cases were to be concerned, I turned out to be right about how the appellate courts would treat the issue on appeals. But I was wrong about the precedential value their decisions, however inappropriately, would come to be accorded.

KUMHO TIRE V. CARMICHAEL[56]

I have always thought *Kumho Tire v. Carmichael* was a great opinion, much more so than *Daubert* itself. As everyone knows, or at least everyone likely ever to read this article, *Kumho Tire* declared that the gatekeeping responsibility of federal trial courts to insure the reliability of expertise under Rule 702 applies with equal force to all proffers, not just to the products of "science."[57] And I venture that that is all most people know about the opinion. But it actually stands for much more.

First, it authoritatively declares that the target of the reliability determination is the exercise of expertise that is involved in the case before the court, not some more general issue concerning the reliability of an area of expertise globally.[58] This aspect of the case is in many ways more important than the clarification in regard to non-science expertise, because it both declares how a court should determine the target issue in the case before it, and by doing so defines the terms by which relevancy of information bearing on reliability is to be judged. In addition, while the opinion does recognize that the "*Daubert* factors" will not be appropriate indices of reliability in regard to every claim of expertise (and therefore that the trial judge needs "flexibility" in making an intelligent assessment of what information rationally could or should bear on reliability in regard to the particular claim of expertise in the case) the *Kumho Tire* opinion likewise makes clear that this "flexibility" is obligatorily in service of arriving at the best set of criteria for actually judging reliability in regard to such expert claims.[59] The case also makes clear that some factors of apparent weight, like whether or not the expert is competently using procedures accepted by others claiming similar expertise, are of little weight when the challenge is based on the lack of reliability of the area as a whole in such applications.[60] Finally, the opinion recognizes that, while the judge must ultimately be affirmatively convinced that the proffered expertise is sufficiently reliable for the

---

Daubert *and* Joiner, 48 Hastings L.J. 969 (1997), especially pages 976–79 (arguing that an "abuse of discretion" standard of review for Rule 702 reliability decisions would remove the courts of appeals from a necessary role in disciplining the lower courts in their approach to reliability decisions).

56. 526 U.S. 137 (1999).

57. *Id.* at 149.

58. *Id.* at 153–54, 155, 156–57.

59. *Id.* at 152–53. See also Justice Scalia's concurrence, *id.* at 158–59 (Scalia, O'Connor & Thomas, JJ., concurring).

60. *Kumho Tire*, 526 U.S. at 151 (majority).

purposes of the law once that issue is put in play, under some circumstances, perhaps especially when the challenge expertise has been admissible by long usage, the opponent cannot simply make a formal objection or demand for a hearing under Rule 702, but must supply the court with sufficient evidence to make a serious and affirmatively warranted issue of the reliability challenge.[61] And such a "price of admission" proffer, a fortiori, ought generally to be task-specific also.

Perhaps it would not be amiss to expand just a bit on the task-specificity requirements of *Kumho Tire*, with a simple illustration from the field of handwriting identification. In *Starzecpyzel*, the only issue involved was whether the signature "Ethyl Brownstone" on two documents was in fact signed by Ethyl Brownstone. This determination, which is often done, as in *Starzecpyzel*, by comparing numerous concededly authentic signatures of the real Ethyl Brownstone ("known signatures") with the "questioned" signature or signatures, is known as "signature authentication." Because of the special circumstances of signature generation applicable to most people (it is a personal symbol, it is important, it is repeated often over long periods of time, etc.) standard document examiner theory holds that such signature authentication is in general among the easiest tasks to perform accurately.[62] There is certainly nothing counterintuitive about this, and there is even some empirical data to support the ability of document examiners to do this task with fairly good success and somewhat better than the ordinary people that make up juries.[63] So it might not be an error to decide that performing the task of signature authentication was sufficiently reliable for the results to be admissible under Rule 702.[64] However, what if the prosecution wanted the document examiner to testify, not only that the signatures were not really signed by Ethyl Brownstone, but that they were signed by Roberta Starzecpyzel? This would present an entirely different task, and a different question. This task (attribution of authorship from an inauthentic signature, or signatures) is looked upon in standard document examiner lore as being a difficult task,[65] especially if the signatures to be attributed were themselves attempts at simulating a genuine Brownstone signature by one with access to such a model. In addition, there is no empirical evidence whatsoever that validates any claim that document examiners can make such attributions reliably.

Now in a "global reliability" approach, a court might look at the empirical data on document examiner skill in regard to signature authentication, and decide that it generally supported some broad skill for doing handwriting analysis tasks in general, and therefore use it to support admitting the testimony as to signature attribution. But

---

61. To use Justice Breyer's words, the reliability of the expert application in the case must be "called sufficiently into question . . . ." *Id.* at 149. "Otherwise, the trial judge would lack the discretionary authority needed both to avoid unnecessary 'reliability' proceedings in ordinary cases where the reliability of an expert's methods is properly taken for granted, and to require appropriate proceedings in the less usual or more complex cases where cause for questioning the expert's reliability arises." *Id.* at 152.

62. *See* Albert S. Osborn, *Questioned Documents* 13 (Lawyers Coop. 1910). "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." *Id.*

63. See Part II of the Appendix, *infra*, especially the summaries of the Kam IV study and the Found and Rogers and Sita, Found, and Rogers studies. *See generally* Park, *supra* n. 40.

64. *See id.*; Risinger, *supra* n. 20, at § 33:39.

65. See quotation from Osborn, *supra* note 62 there set out.

Kumho Tire says that such an approach would be wrong. What we want to know is, can experts do what they claim in the particular application in the case. The evidence about authentication, while not perhaps irrelevant under Rule 401, is not sufficient to establish the examiner's reliability in regard to this "task at hand," that is, attribution of authorship from these two signatures. The task-specific focus of *Kumho Tire* insures that there is good reason to believe that what is actually being done in the case is sufficiently reliable for the purposes of the law, pursuant to Rule 702.

Kumho Tire seemed just like the case I had been waiting for. It incorporated many themes I had embraced for years. Now all that was necessary was for courts to read the opinion and apply it in the cases before them. I settled into my blind to observe.

Well, at least as to my little Cerions, things picked up considerably after *Kumho Tire*, but this seemed to be a total coincidence, since none of the federal decisions in the next nine months seemed influenced by *Kumho Tire* at all. They all cited it, but the citation was perfunctory. The Eleventh Circuit opinion in *United States v. Paul*[66] cited *Kumho Tire* for the "not just science" point, and emphasized passages containing the words "flexibility," "leeway," and "latitude" without noting that the dominant use of the terms in *Kumho Tire* was in regard to making an intelligent and judicious selection of the proper criteria to judge reliability in regard to the claim of expertise in the particular case. Judge Gertner in *United States v. Hines*[67] admitted that she had not fully internalized the approach of *Kumho Tire* when she wrote the opinion that imposed the restriction on the handwriting expert in that case, which she borrowed from Judge Matsch in *McVeigh*, and allowed testimony about similarities and differences, but not the ultimate conclusion concerning attribution of authorship. The Tenth Circuit in *United States v.Battle*[68] cited *Kumho Tire* (once) only for the "not just science" point. The court in *United States v. Santillan*[69] repeatedly cited *Kumho Tire* for the "not just science" point and for the "flexibility" language (even though ultimately settling on a *Hines/McVeigh* approach). The order in *United States v. Brown*[70] is too short to be expected to do much, but, just like the other four cases, it manifested no understanding of the obligation under *Kumho Tire* to define the expert task at issue in the particular case, the "task at hand." So in the nine months after *Kumho Tire*, two circuit courts (*Paul* and *Battle*) issued opinions affirming global admissibility decisions of district courts, based on a global "guild test," while three district courts (*Hines*, *Santillan*, and *Brown*) reached *Hines/McVeigh* results based on some reservations about the global warrant to accept the general reliability of handwriting identification expertise. But not one court did what *Kumho Tire* mandated, or seemed to notice.[71]

66. 175 F.3d 906, 910 (11th Cir. 1999) (Appendix case 7).

67. 55 F. Supp. 2d 62 (D. Mass. 1999) (Appendix case 8).

68. 188 F.3d 519 (table), 1999 WL 596966 (10th Cir. 1999) (Appendix case 9).

69. 1999 WL 1201765 (N.D. Cal. Dec. 3, 1999) (Appendix case 10).

70. No. CR 99-184 ABC (C.D. Cal. Dec. 6, 1999) (Appendix case 11).

71. The reader should note that, under standard theory concerning what is subject to plenary review and what is subject to an abuse of discretion review, if the courts in *Paul* and *Battle* had realized that the lower courts had violated the mandate of *Kumho Tire*, they would generally have had to reverse and remand for a new determination, absent a finding of harmless error, or a failure of retroactivity. This is because explicit failure to follow the legally prescribed approach at all is not discretionary. In fact, one of my set of handwriting cases contains just such a reversal and remand in another context (*Deputy v. Lehman Bros. Inc.*,

So I decided I would write a law review article and straighten things out. Well, I wrote the article. It appeared in the Washington & Lee Law Review in late 2000.[72] I walked the reader through the *Kumho Tire* opinion, with plenty of quotations and annotations, to show how clearly, how inescapably, it declares the "task at hand" approach to be required under Rule 702, and how Justice Breyer goes so far as to "teach by doing," by walking the reader through the proper analysis in regard to the case in front of it instead of remanding for a lower court to perform that function.[73] A lot of my academic friends found it persuasive. But it didn't straighten things out.

In retrospect, it was kind of silly for me to think that something I wrote about a Supreme Court opinion would affect the way judges looked at the opinion. There was growing evidence that judges (well, judges dealing with handwriting reliability challenges, at any rate, and maybe others also) were failing to read the *Kumho Tire* opinion itself, or at least read it very carefully. If they don't read an applicable opinion of the Supreme Court very carefully, why should I think they would even look at anything I might write about it. We live in an age in which judges publicly excoriate the academy for not producing anything of use or value to themselves in performing their judicial tasks. And I know that they are overworked and underpaid and that the scarce resource of judicial time and attention must be husbanded. So once it became clear they weren't paying any attention to the Supreme Court, I realized that I shouldn't be upset that they paid no attention to me.[74]

At any rate, there was still plenty of action among my little Cerions, or as much action as you can expect from snails.

THE POST-*KUMHO* WORLD

Between the beginning of 2000 and the end of 2003, there were 26 federal handwriting reliability decisions, more than twice the number in the preceding seven years since the *Daubert* decision. (In 2002 alone there were eleven.) Of the 26 federal cases, 15 were district court decisions and 11 were appellate court opinions (ten court of

345 F.3d 494 (7th Cir. 2003) (Appendix case 36)), where the circuit reversed the lower court's exclusion of a document examiner and remanded for a new reliability determination because the trial court had explicitly taken personal credibility into account in a way violative of Rule 702. Given the fact that both *Paul* and *Battle* were virtually certain to have been tried before the decision in *Kumho Tire*, and that *Kumho Tire* would almost certainly count as a new construction of Rule 702, perhaps that might have been a basis for not reversing, based on non-retroactivity. But that is not what those courts did.

72. Risinger, *supra* n. 37.

73. This is one of my favorite parts of *Kumho Tire*. Ordinarily, an opinion of the Supreme Court would have gotten to the point at the end of Part II of the *Kumho Tire* opinion, and then remanded for further proceedings to apply the principles just enunciated. However, Justice Breyer apparently believed that an example of "best practices" was necessary to give the lower courts a model to work from, and so he began Part III of the *Kumho Tire* opinion with the following words: "We further explain the way in which a trial judge 'may' consider *Daubert*'s factors by applying these considerations to the case at hand . . . ." 509 U.S. at 153. He then goes on to examine the task specific claims and the evidence for and against them in great detail. This unusual exercise in "teaching by doing" was the basis of Justice Stevens's partial dissent. *Id.* at 159.

74. Not that this stopped me. Along with Professor Denbeaux, I took another, more detailed, run at explaining the joys of the task-specific approach, and how it should work in the context of both scientific and unscientific claims. *See* Denbeaux & Risinger, *supra* n. 45, at 31–60. I have no reason to believe any judge ever read that one either. Federal judges have recently been quoted in the public press as saying that academics no longer write articles that are of any use to them in their work. Perhaps there are some articles that they should actually read before reaching that conclusion.

appeals decisions and one from the Navy/Marine Corps Court of Criminal Appeals). The district court opinions were decidedly mixed in result, with six admitting the proffered handwriting identification testimony (*Richmond*, *Gricco*, *Nadurath*, *Prime*, *Thornton*, and *Jabali*[75]), three imposing the *Hines/McVeigh* limitation (*Rutherford*, *Hidalgo*, and *Oskowitz*[76]), five rejecting the proposed testimony on one ground or another (*Fujii*, *Saelee*, *Brewer*, *Lewis*,[77] and the civil case *Wolf v. Ramsey*[78]), and one indicating that the testimony would have been rejected had the court not ruled that the document to which it pertained was irrelevant (the civil case *Church v. Maryland*). The appellate decisions were as expected—ten of the eleven (*Jolivet*, *Bates*, *Och*, *Elmore*, *Johnson*, *Hernandez*, *Mooney*, *Kehoe*, *Sanders*, and *Crisp*[79]) affirmed the action of the lower courts, and one reversed a lower court exclusion of a handwriting witness on a technical ground unrelated to reliability (*Deputy v. Lehmann Bros.*, a civil case[80]). All of the ten affirmances were in criminal cases, and two were especially easy, since the defense had failed to preserve any claim of error in the admission of the handwriting testimony by properly objecting below (*Jolivet* and *Sanders*).

Meanwhile, in the fall of 2001 the state courts started to come in with decisions post *Kumho*, and by the end of 2003 there were eleven of them, nine criminal and two civil cases (eight appellate and two trial court decisions), and all of them resulted in the admission of the handwriting testimony and, in the case of the appellate decisions, affirmance on appeal.[81] I was happy neither with these results nor with the way they had

---

75. *U.S. v. Richmond*, 2001 WL 1117235 (E.D. La. Sept. 21, 2001) (Appendix case 18); *Gricco*, 2002 WL 746037 (Appendix case 23); *U.S. v. Nadurath*, 2002 WL 1000929 (N.D. Tex. May 14, 2002) (Appendix case 24); *Prime*, 220 F. Supp. 2d 1203 (Appendix case 27); *U.S. v. Thornton*, No. 02-M_9150_01 (D. Kan. Jan. 24, 2003) (Appendix case 31); *Jabali*, 2003 WL 22170595 (Appendix case 35).

76. *U.S. v. Rutherford*, 104 F. Supp. 2d 1190 (D. Neb. 2000) (Appendix case 12); *U.S. v. Hidalgo*, 229 F. Supp. 961 (D. Ariz. 2002) (Appendix case 28); *U.S. v. Oskowitz*, 294 F. Supp. 2d 379 (E.D.N.Y. 2003) (Appendix case 37).

77. *U.S. v. Fujii*, 152 F. Supp. 2d 939 (N.D. Ill. 2000) (Appendix case 14); *U.S. v. Saelee*, 162 F. Supp. 2d 1097 (D. Alaska 2001) (Appendix case 16); *U.S. v. Brewer*, 2002 WL 596396 (N.D. Ill. Apr. 16, 2002) (Appendix case 22); *U.S. v. Lewis*, 220 F. Supp. 2d 548 (S.D.W. Va. 2002) (Appendix case 26).

78. 253 F. Supp. 1323 (N.D. Ga. 2003) (Appendix case 34). *Wolf* is something of a classification problem. One of the plaintiff's handwriting experts was excluded on qualifications grounds, and the other was forbidden from testifying in terms of "100%" certainty, which the court determined left the plaintiff with a failure of proof. The former is not an expertise reliability decision within the terms of this paper, and the latter might be characterized as a *Hines/McVeigh* result, except for the summary judgment that was then entered.

79. *U.S. v. Jolivet*, 225 F.3d 902 (8th Cir. 2000) (Appendix case 13); *U.S. v. Bates*, 240 F.3d 1073 (table), 2000 WL 1835092 (5th Cir. 2000) (Appendix case 15); *U.S. v. Och*, 16 Fed. Appx. 666 (9th Cir. 2001) (unpublished) (Appendix case 17); *U.S. v. Elmore*, 56 M.J. 533 (Navy-Marine Crim. App. 2001) (Appendix case 19); *U.S. v. Johnson*, 30 Fed. Appx. 685 (9th Cir. 2002) (unpublished) (Appendix case 20); *U.S. v. Hernandez*, 42 Fed. Appx. 173 (10th Cir. 2002) (unpublished) (Appendix case 25); *U.S. v. Kehoe*, 310 F.3d 579 (8th Cir. 2002) (Appendix case 29); *U.S. v. Mooney*, 315 F.3d 54 (1st Cir. 2002) (Appendix case 30); *U.S. v. Sanders*, 59 Fed. Appx. 765 (6th Cir. 2003) (unpublished) (Appendix case 32); *Crisp*, 324 F.3d 261 (Appendix case 33).

80. 345 F.3d 494 (Appendix case 36). For the description of *Deputy*, see *supra* note 71, and in Appendix, *infra*.

81. *Basinger v. Commonwealth*, 2000 WL 724037 (Va. App. June 6, 2000) (Appendix case 52); *Est. of Acuff v. O'Linger*, 56 S.W.3d 527 (Tenn. App. 2001) (Appendix case 53); *Taylor v. Abernethy*, 560 S.E.2d 233 (N.C. App. 2002) (Appendix case 54); *Williams v. State*, 60 P.3d 151 (Wyo. 2002) (Appendix case 55); *Spann v. State*, 857 So. 2d 845 (Fla. 2003) (Appendix case 56); *State v. Bradford Jones*, 2003 WL 2159842 (Del. Super. July 2, 2003) (Appendix case 57); *Florence v. Commonwealth*, 120 S.W.3d 699 (Ky. 2003) (Appendix case 58); *Commonwealth v. Murphy*, 797 N.E.2d 394 (Mass. App. 2003) (Appendix 59); *State v. Matthews*, 859 So. 2d 863 (La. App. 2003) (Appendix case 60); *Commonwealth v. Glyman*, 2003 WL 22956121 (Mass.

been reached (dominantly the guild test, globally applied). But I had not expected the states to lead the federal courts. And as to the federal courts, the courts of appeals were doing generally what would be expected under *General Electric v. Joiner*, and a high percentage of the district courts that examined the reliability evidence produced in so-called *Daubert* hearings (which would be better called *Kumho Tire* hearings) had grave misgivings and acted upon them.

I finished up a revision of the Modern Scientific Evidence handwriting Chapter in mid-2004 and turned away to pursue other interests. And when I got back in 2006 my little Cerions were nearly extinct.

POST-2003

Maybe I should have seen it coming. The quality of the opinions generated in all of these cases previously noted was decidedly below excellent. The best of those opinions up to the end of 2003 were arguably Judge Gottschall's opinion in *U.S. v. Fujii*, and Judge Holland's opinion in *U.S. v. Saelee*, because they were the only opinions to operationalize the task-at-hand approach mandated by *Kumho Tire*.[82] And even in that case, both Judge Gottschall and Judge Holland seem to have arrived at the approach by independent invention, because, although they cite *Kumho Tire*, they do not cite it for its requirement of task specificity. Whatever the results of all of the other cases, and whatever their other strengths and weaknesses, every opinion except *Fujii* and *Saelee* approached the reliability decision globally, in derogation of the requirements of *Kumho Tire*. In addition, by the end of 2003 one was just beginning to see (if one were perceptive enough, which I was not) what turned out to be an ominous development: String citation to court of appeals decisions as authority for the proposition that admission of handwriting identification expert testimony was the right thing to do (assuming the paper credentials of the proposed witness).

Anyone who has read this far will see that such use of the court of appeals opinions in the context in which they were rendered is questionable at best. Three (*Ruth*, *Velasquez*, and *Jones*) were pre-*Kumho Tire* and *General Electric v. Joiner*. The other ten criminal cases[83] (*Paul*, *Battle*, *Jolivet*, *Bates*, *Och*, *Elmore*, *Johnson*, *Hernandaez*, *Kehoe*, *Mooney*, *Sanders*, and *Crisp*) were all decided under an abuse of discretion standard, and none of them, predictably, found an abuse of discretion in the admission of

---

Super. Dec. 15, 2003) (Appendix case 61); *People v. Nawi*, 2004 WL 2944016 (Cal. App. 1st Dist. Dec. 21, 2004) (Appendix case 63).

82. Of the opinions that admitted handwriting testimony, only Judge Lasnik's opinion in *Prime* (Appendix case 27) seems both to recognize the "task at hand" requirements of *Kumho Tire* and to make some attempt to take them into account. But, as the analysis in the Appendix makes clear, Judge Lasnik mistakes the actual multiplicity of tasks before the document examiner for a single task, and that undermines his whole analysis, which was based on characterizing the task as one involving a large volume of questioned writing, not many tasks each involving a small amount of writing. Judge Lasnik's opinion is probably rightly viewed as, overall, the best opinion of those admitting the proffered testimony, but it is tortured in two senses of the word, in that he manifests great discomfort with the weaknesses of the evidence for actual reliability and the research done on the issues, and also in that he chops logic to avoid carrying through with the implications of the first part of the opinion, possibly based on his frank assessment that the Supreme Court really didn't mean to reach long used prosecution-proffered expertise.

83. The civil case, *Deputy v. Lehman Bros.*, was *sui generis* for a number of reasons discussed in note 71 *supra* and Appendix, *infra*.

handwriting identification expertise, globally and on the records before them. However, this can hardly stand as authority for the proposition that such admission is either required or advisable. Three of these cases (*Jolivet*, *Bates*, and *Sanders*) were cases raising unpreserved claims. One (*Och*) was resolved exclusively on harmless error grounds. One (*Battle*) manifested some discomfort with the reliability conclusion, and gave an alternative ground of harmless error. Another, the Navy/Marine Corps Court of Criminal Appeals case (*Elmore*), was disposed of largely by citation to the pre-*Kumho Tire* military court's decision in *U.S. v. Ruth*, as if *Ruth* were good authority *after Kumho Tire*. Another (*Johnson*) relies on a *pre-Daubert* opinion as dispositive of the proposition that "handwriting analysis is a science in which expert testimony assists a jury."[84] Another (*Hernandez*) is a short opinion (relying entirely on the expert's credentials and a citation to *Jolivet*) upholding a *Hines/McVeigh* result in the face of defendant's global assertion that handwriting identification testimony should be barred generally. Another (*Kehoe*) deals with the handwriting reliability issue in four sentences giving a summary of the expert's credentials and a citation to *Jolivet*. Another (*Crisp*) was decided 2-1 with a dissenting opinion. Five of these opinions (*Och*, *Bates*, *Johnson*, *Hernandez*, and *Sanders*) are officially "unreported," and all these opinions are contextually weak as precedent, given that they were decided globally (making no attempt at a *Kumho Tire* task-at-hand analysis) and on abuse of discretion grounds. In addition, as just catalogued, most have other weaknesses as authority.

Only two post-*Kumho* court of appeals decisions are without such additional contextual problems—*Paul* and *Mooney*. But the *Paul* opinion, for all its length, actually never addresses the lower court's reliability decision explicitly at all, and only does so implicitly by virtue of its affirmance of the conviction.[85] That leaves *Mooney*. The only reliability issue before the court of appeals was not the admissibility of testimony by the expert, but the refusal to impose a *Hines/McVeigh* limitation on the expert, so that the expert could not give a conclusion actually attributing authorship of the questioned documents. In addition, the rest of the evidence in *Mooney* was so overwhelming that the court could easily have given harmless error as an alternative ground for affirmance. The court does not invoke harmless error, but the weight to be given its resolution of the issue in front of it must be evaluated in that light. But the opinion in *Mooney*, while it does have some problems of analysis, is careful to leave open the possibility that either decision (deciding to impose a *Hines/McVeigh* limitation or deciding not to impose such limitation) would not be an abuse of discretion.[86] So, when actually examined, none of these court of appeals decisions, alone or together, provides a secure basis for their invocation as authority in regard to what a court ought do, even globally, much less under the actual task-specific requirements of *Kumho Tire*.

But the overwhelming problem in using these appellate decisions as proper

84. 30 Fed. Appx. at 688.

85. This oddity was actually noticed by the court in *Lewis*. *See* 220 F. Supp. 2d at 552 (characterizing *Paul* as "concluding summarily that *Daubert* factors do not necessarily apply to admission of handwriting expert, but failing to address any alternative measures of reliability").

86. Rather ironically, the *Mooney* court also notices the *Kumho* task-at-hand language, but turns it to its own uses in distinguishing *Hines*. 315 F.3d at 63 (discussing *Hines*, 55 F. Supp. 2d 62). It is ironic because [illegible] court approach and the rest of the court of appeals opinion in *Mooney* were global.

precedent for what one might call "best practice" Rule 702 decisions is their inevitable skew. While at least one court has noticed the "abuse of discretion" weakness,[87] no court has noticed the skew problem. The skew problem arises because appeals by the government challenging exclusion or limitation of prosecution-proffered expert testimony (including handwriting testimony) are virtually non-existent. That is because, under federal practice with its strong policy against interlocutory appeals, either the criminal defendant is convicted, in which case the issue is moot, or the criminal defendant is acquitted, in which case the government cannot appeal. So the only cases appellate courts see involve situations where the testimony was admitted and the defendant was convicted. What appellate courts would have to say about exclusion or limitation under an abuse of discretion standard is unknown, but it seems likely that, given an appropriate hearing and findings, that result would be most likely be affirmed also.

Well, as I said, if I had noticed the potentials of the questionable invocation of case authority in the few pre-2004 opinions where it was present (*Gricco*[88] and *Nadurath*[89] at the district levels, *Hernandez*[90] and *Crisp*[91] at the appellate level) I might have been less shocked at what I found in 2006-2007. First of all, the last case reflecting a decision to exclude or limit handwriting identification expertise was *Oskowitz*,[92] decided December 10, 2003. From that date to this writing, all 13 federal decisions, trial and appellate, civil or criminal, admitted or affirmed the admission of proffered handwriting identification expertise globally. In addition, the number of decisions diminished rapidly. From the beginning of 2002 until the end of 2003 (two years) there were 18 cases, 11 trial decisions, and 7 appellate decisions. From the beginning of 2004 to the end of 2007 (4 years) there were 13 cases, 6 trial and 7 appellate. In 2007, there was only one decision, a district court decision (*Yagman*[93]). But it is not merely the decline in cases that signals the extinction of my Cerions. The form of the usual opinion became semi-stylized, and highly authority driven, usually built around an inappropriate string-cite to court of appeals decisions, combined with a recitation of the proposed expert's paper credentials, sometimes supplemented by a thoughtless checklisting of the "*Daubert* Factors" using the handwriting guild as the reference class to establish both peer review and general acceptance globally. This general form is reflected in nine opinions (*Prime II*,[94]

87. In the district court opinion in *Prime*, Judge Lasnik recognizes the problem of relying on the precedent of court of appeals decisions decided under *Joiner*. *See Prime I*, 220 F. Supp. 2d at 1208 (Appendix case 27). The irony is that the circuit opinion in *Prime II* affirming Judge Lasnik's decision to allow the handwriting testimony in *Prime I* relied in part on the precedents from the other circuits, 363 F.3d at 1034–35, and that court of appeals opinion in *Prime II* was then used in the same way as virtually binding authority by Judge Wilson in *Yagman*. *See U.S. v. Yagman*, 2007 WL 4409618 at **2–8 (C.D. Cal. May 22, 2007) (Appendix case 50).

88. 2002 WL 746037 (Appendix case 23).

89. 2002 WL 1000929 (Appendix case 24).

90. 42 Fed. Appx. 173 (Appendix case 25).

91. 324 F.3d 261 (Appendix case 33). See also *Prime II*, 363 F.3d at 1032–53 (Appendix case 38), where the string-cite to other appellate decisions was also part of the court's asserted rationale for affirming the lower court's admission of handwriting identification testimony.

92. 294 F. Supp. 2d at 383-84 (Appendix case 37).

93. 2007 WL 4409618 (Appendix case 50).

94. 363 F.3d 1028. The district court and court of appeals opinions in *Prime* both contain rather tortured

Ferguson,[95] *Ojeikere*,[96] *Judson Brown*,[97] *Smith*,[98] *Adeyi*,[99] *Campbell*,[100] *Versace*,[101] and *Yagman*[102]). Three of the other four (*Rutland*,[103] *Mornan*,[104] and *Garza*[105]) were

"*Daubert* factor" analyses, but the circuit does not even do justice to the job done below. For instance, the entirety of what it says about "general acceptance" is that the district court "recognized the broad acceptance of handwriting analysis and specifically its use by such law enforcement agencies as the CIA, FBI, and the United States Postal Inspection Service." *Id.* at 1034. The string-cite to six other circuit opinions follows in the next paragraph. *Id.* at 1034–35.

95. *U.S. v. Ferguson*, 2004 WL 5345480 (S.D. Ohio July 30, 2004) (Appendix case 40). *Ferguson*'s reliability rationale is entirely authority-driven. It starts with the string-cite, *id.* at *7, then concentrates on the authority of (of all cases) *Jones*. *Id.* at **7–8. This is understandable enough, I suppose, since *Jones* is a Sixth Circuit case and Ohio is in the Sixth Circuit.

96. 2005 WL 425492 (Appendix case 41). *Ojeikere* could have been omitted from the list of reliability decisions on the ground that the court never actually reached or formally resolved the reliability issue. But considering the fact that Judge Koeltl appears not to think a reliability hearing was even necessary, and seems to have scheduled one only because the prosecution favored one (likely to allay a possible issue on appeal), see *id.* at *4, it seems fair enough to include it, given the notation that "[v]arious courts of appeal have uniformly affirmed the admissibility of expert handwriting analysis over *Daubert* objections." *Id.* (But to be fair, also noting that "some district courts have . . . excluded or limited such testimony"), and the fact that the *Ojeikere* opinion is the only one to wholeheartedly embrace the government's assertion that the Advisory Committee's citation to *Jones* in the note to amended Rule 702 "specifically contemplate[s] the admissibility of expert handwriting analysis based on the experience of a proposed examiner . . . ." *Id.* at *3. See the discussion of that position in *supra* note 45.

97. *U.S. v. Brown*, 152 Fed. Appx. 59 (2d Cir. 2005) (unpublished) (Appendix case 43). The handwriting aspect is disposed of in one paragraph centered on the string-cite. *Id.* at 62. The decision is backed up by a second paragraph invoking harmless error. *Id.* at 63.

98. 153 Fed. Appx. 187 (4th Cir. 2005) (unpublished) (Appendix case 44). *Smith* is arguably the most extreme example of a precedent-driven approach. The district court admitted the testimony of the expert without a hearing. The circuit court decided that the previous decision in *Crisp*, also a Fourth Circuit case, justified the decision not to hold a hearing. *Id.* at 190. Thus *Crisp* was interpreted as laying handwriting identification reliability issues to rest globally, or at least as globally authorizing refusal of any Rule 702 hearing in regard to such claimed expertise.

99. *U.S. v. Adeyi*, 165 Fed. Appx. 944 (2d Cir. 2006) (unpublished) (Appendix case 45). Like *Gonzales*, *Jolivet*, *Sanders*, and *Mornan*, this case deals with unpreserved error, since the handwriting evidence was not properly objected to at trial. On that basis I would not ordinarily include it in a list of the precedent-driven opinions actually disposing of reliability challenges. However, the opinion does contain a footnote with the usual string-cite strongly suggesting, though not formally holding of course, that such testimony is always admissible. *See id.* at 946 n. 1. The defense later raised the issue of ineffective assistance of counsel in a habeas petition, and the district judge rejected that claim based on the fact that the testimony would have been admissible even if objected to, citing the circuit court footnote, and otherwise basing his (global) decision entirely on judicial authority. *See Adeyi v. U.S.*, 2007 WL 203962 at **4–5 (E.D.N.Y. Jan. 24, 2007). For this reason, the whole *Adeyi* package is here included in the list of cases resolving reliability issues on precedent grounds.

100. *U.S. v. Campbell*, 2006 WL 346446 (N.D. Ga. Feb. 13, 2006) (Appendix case 46). This short opinion is entirely based on precedent. *See id.* at *3.

101. *A.V. By Versace, Inc. v. Gianni Versace*, 446 F. Supp. 2d 252 (S.D.N.Y. 2006) (Appendix case 48). Judge Leisure deals with the reliability challenge in a footnote, on the assertion that his decision to admit the expert was made "[m]indful of the *Daubert* factors," plus a circuit court string-cite. *Id.* at 268 n. 14.

102. 2007 WL 4409618. The opinion in *Yagman* is fairly lengthy, but it is largely devoted to importing the general reasoning in *Prime II* (a decision of the Ninth Circuit in which the district is located) into the opinion in *Yagman*. It also concludes with a circuit decision string-cite, and while it recognizes that some district courts have rejected or limited handwriting testimony, it concludes that the persuasiveness of those decisions "is extremely limited in light of more recent circuit court decisions permitting expert handwriting testimony." *Id.* at *9.

103. *U.S. v. Rutland*, 372 F.3d 543 (3d Cir. 2004) (Appendix case 39). *Rutland* involved an objection that the proffered expert was too qualified, that is, that his qualifications were so impressive relative to the reliability of the area that his testimony would inevitably be prejudicially overvalued. In rejecting this argument, the court noted that such a ruling would force litigants to search for apparently less well qualified experts, and make judgments about whether they were hiring someone who crossed an elusive line of being too impressive. *See id.* at 546. I was of two minds about whether to include the case on the list of cases dealing with reliability issues, but finally decided to err on the side of inclusion.

sui generis, for reasons explained in the footnotes and Appendix, and the fourth (*Truman Arnold*[106]) was so short and conclusory that the district judge apparently felt that even a string-cite was unnecessary.

So it has come to this. I am packing my bags and leaving the island unless and until some future event (a decision of the Supreme Court, perhaps, explaining that they really meant what they said in *Kumho Tire*) creates a more conducive climate for the healthy and rational examination of the reliability problems of various aspects of handwriting identification expert testimony.[107]

CLEAN-UPS AND TAKE-AWAYS

What are we to make of the evolution of judicial treatment of handwriting expertise? The decisions, viewed in the light of *Kumho Tire*, have generally been poor performances, yet the judges involved are judges who, in other contexts, in regard to other kinds of issues, have generated excellent opinions, I am sure.[108]

One way to view what has happened is to see it as the courts turning a blind eye to a mandate of the Supreme Court to approach expert reliability issues task-specifically.

---

104. *U.S. v. Mornan*, 413 F.3d 372 (3d Cir. 2005) (Appendix case 42). *Mornan*, like *Gonzales*, *Jolivet*, and *Sanders*, deals with an unpreserved objection to handwriting testimony. The main issue raised on appeal seems to have been that it was plain error to let the document examiner testify in qualified language, which the court properly rejected. *Id.* at 381. No actually reliability issues are explicitly addressed, although by implication the court finds the testimony to be reliable enough that its admission was not plain error, hardly a surprising conclusion.

105. *U.S. v. Garza*, 448 F.3d 294 (5th Cir. 2006) (Appendix case 47). In one way, *Garza* takes the anomaly prize in this set of cases, and in another, it is all of a piece with general judicial predelictions. *Garza* does deal with a reliability challenge, all right, and it confirms the rejection of proffered expert testimony, but the handwriting expert involved was proffered by the defense. In addition, the ground for rejection was based on the fact that the expert did her comparison from photocopies. Generally, this has never been recognized as a proper basis for objection in and of itself. A typical opinion accepting the use of photocopies (by a prosecution expert) is *Nawi*, 2004 WL 2944016 at *19 (Appendix case 63). More importantly, so far as I know, there has never been another case where the sole basis for the exclusion of the testimony of a handwriting expert was the use of copies for comparison purposes. *See* Appendix, *infra* n. 523. In addition, and most importantly, the defendant was in no position to obtain originals. This case is probably better seen as reinforcing the conclusions that are sometimes reached concerning an institutional pro-prosecution bias in the handling of expert reliability issues than anything else. *See* Risinger, *supra* n. 17, at 135–43 (examining inference of pro-prosecution bias flowing from the results of various ways in which different kinds of expertise have been dealt with when subject to *Daubert* challenges).

106. *Truman Arnold Cos. v. Green*, 2006 WL 5153151 (E.D. Tex. Sept. 28, 2006) (Appendix case 49).

107. One potential objection to my systemic conclusion is that it is notoriously difficult to reason with confidence from the evidence of written opinions to the actual general practice in that vast majority of cases that leave no accessible written traces. Certainly some judges are likely to continue imposing *Hines/McVeigh* restrictions if they have already done so. Judge Gertner, for instance, who of course authored the opinion in *Hines*, has extended the *Hines/McVeigh* approach to other forms of claimed forensic identification expertise. *See e.g. U.S. v. Green*, 405 F. Supp. 2d 104 (D. Mass. 2005) (applying *Hines/McVeigh* to firearms identification/toolmark expertise in regard to shell casings). There is the theoretical possibility, therefore, that there is a large reservoir of cases rejecting or restricting prosecution handwriting evidence under Rule 702 that simply has not generated any available writing. However, this seems vanishingly unlikely. Given my various connections to other interested academics, and to sources of information reflecting the general responses and opinions of the document examiner community, if there were any significant number of such cases, I am confident that I would have heard.

108. Just as an example, consider the precise, lawyerly, and generally excellent exposition on the issue of tracing funds in a money laundering case by Judge Wilson in the only "formally" published opinion (in the sense that Judge Wilson actually submitted the opinion to the Federal Supplement) in the *Yagman* case, 502 F. Supp. 2d 1084.

In this area, at least,[109] the lower courts seem unable to come to grips with what is required, though *Kumho Tire* is clear enough, and, as previously noted, Justice Breyer, in an exercise of "teaching by doing," even went to the unusual extreme of walking through the expected process in regard to the issue in that case, rather than simply remanding it for a lower court to conduct the required analysis in a manner "not inconsistent with the court's decision."[110] I believe that the lower courts have instinctively understood that such an approach would be a lot of work, would involve them in complex issues of empirical evaluation of claims beyond their comfort level if not their competence, and most importantly after all that work, would not yield results with much dispositive carry-over application to the next case arising, except in regard to narrow task issues that they were not even sure how to formulate.[111] What they instinctively wanted was a system that yielded broad dispositive precedent, thus obviating the need for repetitively examining an area (like handwriting) in order to map out the real reliability contours of claimed expertise.[112] In order to achieve that, they had to fail to see, never mind understand, the task-at-hand mandate of *Kumho Tire*, because only then could the legal questions be formulated and treated globally enough to be domesticated to a precedent-based system. And so they did. Then, when decisions, especially appellate decisions, were rendered, they were treated as strong authority in resolving the global issues of reliability ("Is handwriting expertise reliable enough to be admitted?" "Does handwriting expertise pass muster under Rule 702?"), even though the context of those decisions deprived them of any binding authority or claim to significant precedential weight under the official doctrines and assumptions of the precedent system. By doing this, the courts turned a process of reliability evaluation which the Supreme Court seems clearly to have envisioned as being empirical, data driven, and specific to the task being performed by the expert in the case at bar, into a single global issue resolvable at a swoop by reference to precedent for all the tasks undertaken by "handwriting experts."

Were I to stop here, I might rightly be criticized for unrealistically making it appear that judges have been consciously and willfully disregarding the requirements of the law just to make their lives easier. I admit that the judicial choices that have led to

109. There is quite a bit of reason to believe that what I have said here about handwriting challenges applies equally to most if not all areas of traditional forensic identification expertise. The global framing of issues, the distorted and mechanical approach to "*Daubert* factors" and the "appeal to precedent" has been even more uniform in cases involving challenges to fingerprint identification than in the handwriting area. *See* Jennifer L. Mnookin, *The Validity of Latent Fingerprint Identification: Confessions of a Fingerprint Moderate*, ___ L. Probability & Risk ___ (forthcoming 2007). The same is certainly true in regard to toolmarks. See Judge Gertner's lament in *Green*, 405 F. Supp. 2d at 122–23. Indeed, I know of no area of traditional forensic identification science in which the pattern is not true in general.

110. *See supra* n. 73 and accompanying text.

111. This is consistent with what I heard at a conference some years ago from a federal district court judge, whose response to me (on the topic of *Kumho Tire* and forensic science) was approximately, "If the Supreme Court thinks I'm going to have a hearing on every little objection to the reliability of this stuff, they're crazy."

112. This is not a trivial issue, and it is reasonable for the courts to have this concern. The question is, how does a court define a case-specific task narrowly enough to meet the policies behind *Kumho Tire* but broadly enough to yield at least some stare decisis carry-over effect. I believe that there are ways of systematically resolving the problem if it is faced explicitly, which Professor Denbeaux and I have set out in detail in Denbeaux and Risinger, *supra* note 45, at 35–55, particularly pages 54–55. *See also* Park, *supra* n. 40 (recognizing the problems of defining the specific task sufficiently broadly to have some appropriate carry-over effect).

the current state of affairs are not likely, in the main, to have been based on such a cold blooded conscious choice. In that regard, what I have just set out is something of a "just so" story. It maps on to what we can observe in the world fairly well in a general sort of way, but the actual forces behind the way the cases have worked their way out are likely to have been more complex. This I readily concede. In fact, I believe there are a number of important identifiable factors beyond the desire for global finality that have contributed to the way things have evolved.

First on my list would be lawyers, most particularly, criminal defense lawyers. These issues have been raised predominantly in criminal cases. Of the 50 federal cases in the Appendix, 46 have been criminal cases. Criminal defense attorneys, as a group, tend to be overworked and have fewer resources than the prosecution. Because they are always "putting out fires" in regard to various cases, issues in any particular case are often seen late and processed under intense time pressure. Formulating a task-at-hand attack requires understanding *Kumho Tire*, understanding the expert claim in the case and where it fits with other kinds of expert claims in the area, and assembling the sources of empirical data that might be said to bear on the task-specific reliability. This is a tall order to fill, on the fly and under pressure, for people without any particular background in the area of expertise at issue. And so global objections get made, which invite judges to undertake global consideration and resolution. In the long run, I believe criminal defense attorneys have contributed heavily to the system's poor performance in this area under the requirements of *Kumho Tire*. I cannot imagine that one could select 50 reliability attacks in civil cases in federal courts in the same time period dealing with an issue like causation in toxic torts, and not see a much more detailed and sophisticated airing of the issues than appears to have been the case in regard to handwriting expertise.[113]

I might also include prosecutors, since they certainly haven't helped. However, in an adversary system, when objections are made to their witnesses, they are not expected to help. If they package arrant nonsense like the Rule 702(b) fallacy into briefs, it isn't necessarily on them that they can get some judges to bite. I am sure that the string-cite approach is pushed in prosecution briefs. But it is incumbent on the judges, and the defense attorneys, to figure out the fallacies of and reject such ultimately untenable arguments, more than it is, alas, on the prosecutors for making them in "our adversary system."[114]

Next on my list would be certain judicial tendencies besides the tendency generally to desire a precedent-based global resolution to issues of expert reliability. First, there is some reason to believe that judges as a group are resistant to rejecting prosecution proffers of expert testimony. One could argue about the various reasons this might be the case, having to do with judicial selection criteria and the like, but I believe that when the issue is expertise that has been admitted without question for generations, and which

113. For an example of one excellent opinion written after what was obviously a sophisticated and complete exposition of the issues on both sides, see *In re Ephedra Prods. Liab. Litig.*, 393 F. Supp. 2d 181 (S.D.N.Y 2005) (Rakoff, J.).

114. Some of these issues are close ones, and perhaps I am being too easy on prosecutors. *See* Paul C. Giannelli & Kevin C. McMunigal, *Prosecutors, Ethics, and Expert Witnesses*, 75 Ford. L. Rev. 1493 (2007).

has played a role in convicting many people (often in cases tried in front of those same judges in the past), and which the many judges who have been prosecutors before ascending to the bench have used in trying cases and convicting defendants, then the resistance becomes intense.[115] I suspect that here, many judges start to regard the reliability attacks as some sort of trick, something that just can't be serious, and that this attitude also drives the way in which they conceive and consider the issues, in order to be done with them once and for all. Certainly something like this tendency seems to be borne out by the results in the state cases. As earlier noted, the first state reliability decision of the post-*Starzecpyzel* era, at any rate, appeared in late 1997 (*Cochran*). The next was decided in late 2002, and since that time, there have been 15 more, for a total of 17 (15 criminal, two civil[116]). Of these cases, which of course are not directly subject to *Kumho Tire*, all have globally admitted handwriting expertise. Most strikingly, at least three appellate courts have established absolute rules essentially instructing their courts not to entertain any reliability challenges to handwriting expertise at all.[117]

And look at the opportunity we have missed to obtain intelligent judicial examination of these issues. There have been 66 cases represented in the Appendix list[118] of reliability challenges. Of the available documents, eight (all federal) reveal such sparse facts that the task at issue cannot be determined at all (*Richmond*, *Nadurath*, *Kehoe*, *Jabali*, *Oskowitz*, *Judson Brown*, *Campbell*, and *Truman Arnold*). An analysis of the remaining 58 cases, state and federal, reveals at least the following task issues that ought to have been examined for reliability under proper standards actually relevant to determining reliability: (1) signature authentication by comparison with authentic signatures; (2) determination of forgery by tracing; (3) determination of disguise attempts in giving demand exemplars; (4) attribution of authorship where there is extensive writing in a single questioned writing, and there is extensive known writing; (5) attribution of authorship when the questioned writing consists of many short writings such as entries on forms and the like, which may or may not be related; (6) attribution of authorship from multiple inauthentic signatures which may or may not be the product of some attempted disguise; (7) attribution of authorship from a single inauthentic signature which may or may not be the product of some attempted disguise; (8) the use of photocopies as exemplars; and (9) the use of photographic copies as the source of unknown writings for attribution.

I believe that, had the courts actually examined these issues, they would almost certainly have come to different results than they did in some cases. In addition, such

115. "There is almost no expert testimony so threadbare that it will not be admitted if it comes to a criminal proceeding under the banner of forensic science." Jane Campbell Moriarty & Michael J. Saks, *Forensic Science: Grand Goals, Tragic Flaws, and Judicial Gatekeeping*, 44 Judges' J. 16 (2005).

116. All are collected in the Appendix, cases 51–67.

117. The most clearly categorical is the Supreme Court of Florida in *Spann*. Two other decisions that will almost certainly be given that effect are the opinion of the Wyoming Supreme Court in *Williams*, and that of the North Carolina Court of Appeals in *Taylor v. Abernethy* (which, though a civil case, may be the most explicit on the point). Also capable of that construction is the decision of the Virginia Court of Appeals in *Basinger*. Of course the opinions leave open the possibility of challenging the reliability of the particular expert, that is, their lack of training and experience, as a potential basis for exclusion.

118. There are 67 decisions, but *Prime* is a single case represented by two decisions, making a total of 66 cases.

scrutiny would have provided an invaluable incentive to the document examiner community to participate in real research that could reveal the reliability contours of their claimed expertise. But, barring some miracle, this opportunity has passed. The system has treated these issues as an oyster treats a grain of sand, an irritant to be walled off and forgotten. Game over.

CONCLUSION

I want to use this concluding section to end on a slightly more optimistic note. I have just said that, when it comes to attempting to use the courts as a vehicle for insuring that prosecution-proffered forensic science expertise is reliable, the game is over. But perhaps that is a bit of an exaggeration. While I still believe that, based on the available evidence, the statement is most likely true systemically, there are still likely to be judges who are receptive to arguments invoking the clear mandate of *Kumho Tire* when it is done clearly and with skill. In addition, there are things that could start a new game. I have already mentioned the possibility of a new decision by the Supreme Court, but the report of the congressionally-mandated National Academy of Sciences/National Research Council Committee on the Needs of Forensic Science (which is expected to be issued before this article reaches print) could have that effect, depending on what it says. Finally, perhaps the courts have already had an effect in helping to energize the forces within the forensic science community who believe in research and reform, and want to lessen the impact of partisanship in their practice,[119] and strengthen the epistemic warrant and scientific acceptability of their product.[120] I think that it is on such fronts that efforts can be best spent, and that is where I have increasingly turned my own focus. So maybe I am just changing addresses. Maybe I have not really said goodbye at all.

119. *See e.g.* Krane et al., *Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation*, 53 J. Forensic Sci. ___ (forthcoming 2008).

120. For a good treatment of many of the things that need to be done, see Paul C. Giannelli, *Wrongful Convictions and Forensic Science: The Need to Regulate Crime Labs*, 86 N.C. L. Rev. 163 (2007).