Brave New "Post-*Daubert* World" — A Reply to Professor Moenssens

D. Michael Risinger[*]
Mark P. Denbeaux[**]
Michael J. Saks[***]

[*] Professor of Law, Seton Hall University School of Law. B.A., Yale University, 1966; J.D., Harvard University, 1969.

[**] Professor of Law, Seton Hall University School of Law. B.A., College of Wooster, 1965; J.D., New York University, 1968.

[***] Edward F. Howrey Distinguished Professor of Law and Professor of Psychology, University of Iowa. B.A., B.S., Penn State University, 1969; M.A., Ohio State University, 1972; Ph.D., Ohio State University, 1975; M.S.L., Yale University, 1989.

The authors are listed in the same order as they were in the main article dealt with by Professor Moenssens; no inference of relative responsibility should be drawn. We would like to thank our research assistant, Jennifer Gaeta, whose skill and dedication made the speed of this response possible.

INTRODUCTION

In Andre Moenssens' recent article *Handwriting Identification Evidence in the Post-*Daubert *World,*[1] certain articles and other activities of the present authors were a primary focus of criticism.[2] Much of that criticism was factually inaccurate and, we believe, unjustifiably intemperate. As such, it requires a response. However, it is important to separate the *ad hominem* aspects of Professor Moenssens' piece from the ideas expressed there and to evaluate the latter on their merits. In an attempt to maintain such separation to the extent productively possible, we have structured this article as follows: Part I will be confined to the "merits" critique just referred to. When it is necessary to deal with Professor Moenssens' views about us in Part I, we have attempted to be as brief and moderate as fairness both to Professor Moenssens and ourselves will allow. In Part II, we have responded in detail to Professor Moenssens' charges against us, and have given our evaluation of his analysis and the care with which he has made those charges.

PART I

A. *Moenssens' main theses and criticisms*

We believe the main outlines of Professor Moenssens' position may be fairly summarized as follows:

- Many people may be called to the witness stand by lawyers as asserted experts to offer opinions concerning the genuineness of a signature or the authorship of some questioned writing.[3]
- Many of those who are willing to testify have no such expertise.[4]

---

1 66 UMKC L. REV. 251 (1998) [hereinafter Moenssens, *Post-*Daubert *World*].

2 Specifically, 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) [hereinafter Risinger et al., *Exorcism*] (adopting Moenssens' usage) and D. Michael Risinger & Michael J. Saks, *Science and Nonscience in the Courts:* Daubert *Meets Handwriting Identification Expertise,* 82 IOWA L. REV. 21 (1996) [hereinafter, Risinger & Saks, *Science and Nonscience*]. Although Professor Denbeaux was not a co-author of *Science and Nonscience,* he has read it and fully agrees with and adopts its positions. We have therefore generally used the term "we" in referring to positions taken in both articles in order to avoid unnecessarily awkward qualifications of diction.

3 *See* Moenssens, *Post-*Daubert *World, supra* note 1, at 252-53.

4 *See id.* at 259.

- However, within the universe of people willing to so testify, there is an identifiable subset who are sufficiently skilled and dependable that there ought to be no question of their qualification to testify.[5]
- That subset consists of persons who have been trained in "Osbornian"[6] principles of handwriting identification and have had sufficient clinical experience in applying those principles in practice.
- It was wrong of Saks, Denbeaux, and Risinger in their 1989 *University of Pennsylvania Law Review* article entitled *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification "Expertise"* (*Exorcism*)[7] not to distinguish explicitly this subset and not to recognize that their critiques, to the extent they had any validity at all, did not apply to this subset.
- In addition, if there were ever any tenable skepticism concerning properly trained Osbornian practitioners, it has been totally eliminated by the study reported by Kam et al.[8] in 1997.[9]
- Indeed, these practitioners are worthy of the name "scientist,"[10] and it was wrong of the Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals Inc.*[11] to set up a false dichotomy between "science" and non-scientific expertise, based on a Newtonian or Popperian view of "science," which was narrow and outmoded when *Daubert* invoked it.[12]
- While it is true that handwriting identification and many other forensic science disciplines, as well as many clinical disciplines outside of forensic science, may not be able to meet the kind of reliability criteria set out in *Daubert*, there is another model for establishing reliability of such a clinical sci-

---

[5] *See id.* at 257-58, 310-29.

[6] That is, methods pioneered by Albert Osborn in the first edition of his book. *See generally* ALBERT S. OSBORN, QUESTIONED DOCUMENTS (1910). For an overview of the Osbornian approach see *Science and Nonscience, supra* note 2, at 67-74.

[7] Risinger et al., *Exorcism, supra* note 2.

[8] Moshe Kam, Gabriel Fielding & Robert Conn, *Writer Identification by Professional Document Examiners*, 42 J. FORENSIC SCI. 778 (1997) [hereinafter Kam II], to distinguish it from the earlier study, Moshe Kam, Joseph Welstein, & Robert Conn, *Proficiency of Professional Document Examiners in Writer Identification*, 39 J. FORENSIC SCI. 5 (1994), [hereinafter Kam I].

[9] *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 311-12.

[10] *See, e.g., id.* at 311-12.

[11] 509 U.S. 579 (1993).

[12] *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 274-85.

ence which is much more appropriate to its situation and realities.[13]

Professor Moenssens then goes on to describe his new proposed model in detail.[14] We will examine the Moenssens model later, but fundamentally it is a process check[15] that depends upon the existence of what is essentially a guild.

Professor Moenssens seems to have a number of goals in his article: (1) to alert the reader to doctrinal differences and differences in training and experience among people called as handwriting experts in American courts; (2) to establish the superiority of one doctrine, one method of training, and one group; (3) to recapture the label of "science," with its attendant rhetorical power, for handwriting identification, at least when undertaken by the right group; and (4) to guard against any residual risk that such disciplines might be found too unreliable to be the subject of testimony, first by trashing the credibility of critics by any available means and, second, by proposing a standard of "reliability" tailored both to insure admissibility and to establish his favored group as the monopoly guild. Finally, the underlying hope that appears to be reflected in the Moenssens article is that, if the rest of these goals are accomplished, the practical effect will be to ensure that fact finders are not provided with any skeptical interpretation of either the methodology or the available validity data (or lack thereof) from any person not certified by the guild.

It is difficult to know how many people currently exist who give or have given expert testimony on handwriting identification issues, whether only once or hundreds of times. Magnitude guesstimates have been given ranging from 1200 to around 5500.[16] It is undoubt-

---

[13] *See id.* at 294-97.

[14] *See id.* at 291-94.

[15] By process check, we mean any legal device that claims or hopes to impose substantive limits on something indirectly as a byproduct of requiring certain procedures. Or, as Moenssens himself says, "The issue ought not to be whether opinions by practitioners in a particular forensic science profession can be quantified statistically in terms of error rates, but whether the opinion of a practitioner was arrived at by following a 'scientific method' of comparison." *Id.* at 320; *see also infra* notes 114-17 and accompanying text.

[16] The lower estimate comes from the testimony of Robert Muehlberger in the *Daubert* hearing in *United States v. Martin. See* Transcript of Proceedings at 137, United States v. Martin, No. 1:96-CR-287-JEC (N.D. Ga. Jan. 22, 1997) (unreported copy on file with authors) [hereinafter *Martin* Transcript]. At that time Muehlberger was chairman of the Questioned Documents section of the American Academy of Forensic Science (AAFS) and vice president of the American Society of Questioned Document Examiners (ASQDE). *See id.* at 133. The higher estimate comes from Moenssens himself, when he says, in *Post*-Daubert *World,* "[T]he total

edly true that there is a great variety of theory, methodology, training, and experience among people who are willing (and are permitted) to testify to handwriting identification issues as experts. It is also true that such experts can profitably be divided into two groups: "orthodox Osbornians" (practitioners trained in the methods pioneered and distilled by Albert S. Osborn in 1910),[17] and the more numerous graphologically influenced practitioners (who may or may not use some or all Osbornian methods, but who claim that graphological study,[18] whether effective in divining personality or not, improves one's ability to identify handwriting).[19]

Beyond these purely descriptive details, however, things that are "undoubtedly true" become scarcer. In the *Exorcism* article, we approached the question globally by asking what data then existed to indicate that handwriting experts in general, or that any of them in particular, could do some or all of the things they claim to be able to do and are willing to testify to. It was not that we were not aware of potential subgroupings, or the possibility of variations among subgroups. It was simply that until the extant data were identified globally, such subgroup questions could not be usefully addressed. We then looked for such data and found that either a tiny amount existed, or none, depending on how one regards the set of proficiency

---

number of persons deemed competent to act as questioned document examiners by the two major professional societies is less than 500" and that the graphologists who testify as to handwriting identification "possibly outnumber them more than 10 to 1." Moenssens, *Post-*Daubert *World*, *supra* note 1, at 265 n.43.

[17] *See generally* OSBORN, *supra* note 6. The term "orthodox Osbornian" or simply "Osbornian" was originally coined in *Science and Nonscience*, *supra* note 2, and then adopted by Moenssens. *See* Moenssens, *Post-*Daubert *World*, *supra* note 1, at 257.

[18] Graphology is generally understood to be the attempt to determine personality traits from examination of handwriting. *See* WEBSTER'S NEW COLLEGIATE DICTIONARY 501 (1977).

[19] Curiously, the orthodox Osbornians cannot agree among themselves about how many people are truly qualified as handwriting identification experts, even in the face of the certification program of the American Board of Forensic Document Examination (ABFDE) (the credentialing arm sponsored jointly by the ASQDE and the AAFS Document Section and the Canadian Society of Forensic Science). *See* Moenssens, *Post-*Daubert *World*, *supra* note 1, at 257 n.15. This is because many government agencies, most notably the FBI, do not require that their document examiners obtain certification beyond in-house training, and the ASQDE is unwilling to take the position that certification is a sine qua non of competence, especially since all the older members of the ASQDE were grandfathered into certification without testing when the ABFDE was established in 1977. Estimates of competent Osbornian Questioned Document Examiners (QDEs) range from 180 (according to FBI document section chief Ronald Furgerson, reported in D. FISHER, HARD EVIDENCE, 196 (1995)), to "not more than 300-500," (Moenssens, *Post-*Daubert *World*, *supra* note 1, at 270), to "about 500." (Muehlberger testimony, *Martin* transcript at 138, *supra* note 16).

test results we have collectively referred to as the Forensic Sciences Foundation (FSF) studies.[20]

Moenssens makes a number of explicit substantive criticisms of our approach repeatedly and at length: (1) that our search for data was not thorough,[21] (2) that we erred in believing that quantified data on handwriting identification ought to exist,[22] and (3) that we erred in concentrating on quantified data and excluding anecdotal case studies as "data."[23] Finally, Moenssens claims that we misinterpreted or misused the data we did find,[24] which he asserts are not in fact relevant data, but are essentially meaningless. We now must turn to these substantive criticisms.

1. The claim that our search for data was not thorough

Professor Moenssens' arguments concerning the sloppiness of our search for data are expansions on similar charges made by the Galbraiths[25] between 1989 and 1995, which we dealt with in our 1997 article *Science and Nonscience in the Courts:* Daubert *Meets Handwriting Identification Expertise* (*Science and Nonscience*)[26] (a fact Moenssens fails to note). Moenssens' contribution to the debate consists mainly of condemnatory adjectives, and the addition of Dutch to the list of foreign languages in which we failed to search for publications. Moenssens' whole point strikes us as a bit disingenuous, since the published

---

[20] In 1975, under a grant from the Law Enforcement Assistance Administration, the Forensic Sciences Foundation (FSF) set out to create proficiency tests for forensic expert specialties, among them handwriting identification. The results of the 1975 pilot test were later reported in JOSEPH L. PETERSON, ELLEN L. FABRICANT, & KENNETH S. FIELD, CRIME LABORATORY PROFICIENCY TESTING RESEARCH PROGRAM, FINAL REPORT (1978). A permanent yearly testing program was begun in 1978 and a handwriting component was added in 1984. *See* Risinger et al., *Exorcism, supra* note 2, at 740. These tests were "operated and maintained by Collaborative Testing Services, Inc." (a contract consultant in test design) "with assistance provided by the Forensic Sciences Foundation," and technical supervision "by the Proficiency Advisory Committee (PAC), a committee of the American Society of Crime Laboratory Directors." CRIME LABORATORY PROFICIENCY TESTING PROGRAM, QUESTIONED DOCUMENTS ANALYSIS, REPORT NO. 89-5. These tests have come to be referred to collectively as the "FSF studies."

[21] *See* Moenssens, *Post-*Daubert *World, supra* note 1, at 300-02.

[22] *See, e.g., id.* at 306.

[23] *See id.* at 301-02.

[24] *See, e.g., id.* at 302 n.228.

[25] Oliver Galbraith III, Craig S. Galbraith, & Nanette G. Galbraith, *The Principle of the "Drunkard's Search" as a Proxy for Scientific Analysis: The Misuse of Handwriting Test Data in a Law Review Article,* 1 INT'L J. FORENSIC DOCUMENT EXAMINERS 7 (1995). Earlier versions of this article were circulated in photocopied form at various document examiner meetings from 1989 to its formal publication in 1995.

[26] *See* Risinger & Saks, *Science and Nonscience, supra* note 2, at 58.

description of our literature search in *Exorcism* explicitly points out that we searched the Journal of Criminal Law and Criminology, which, we further explicitly noted, contains abstracts of significant articles published throughout the world.[27] Be that as it may, the important thing is that neither the Galbraiths nor Moenssens, nor anyone else, has in ten years of highly motivated looking, found a single study generating data on the issue of handwriting identification validity that existed at the time of *Exorcism*'s publication and was not dealt with in that article.[28] So either we were very lucky, or our search strategy was better than Moenssens is willing to credit.[29]

2. The claim that we erred in expecting to find quantified validity data

Professor Moenssens' position on quantified validity data is startling. He opposes "a drastic rule limiting expert opinion evidence only to testimony that can be shown to be reliable by reference to

---

[27] *See* Risinger et al., *Exorcism, supra* note 2, at 784.

[28] That no such studies exist has been attested to by others who have been motivated to look, found nothing, and said so: "[T]here certainly has been a shortage of studies comparing handwriting identification expertise with non-expertise . . . ." Galbraith et al., *supra* note 25, at 7 n.7. There is an "admittedly sparse history of carefully controlled empirical studies." *Id.* at 7.

> If there is a conclusion that can be drawn from the comprehensive literature search performed by Risinger et al. . . . it is that good tests for determining the existence or nonexistence of handwriting expertise need to be devised and that there is a lamentable lack of empirical evidence about the subject in the forensic literature.

Kam I, *supra* note 8, at 7.

The issue of QDE ability is "characterized by acute lack of empirical evidence on the proficiency of document examiners," and that as a result of this deficiency "it is widely agreed that testing of professional document examiners and acquiring data on their abilities . . . are necessary." Kam II, *supra* note 8, at 778.

[29] After all the sound and fury about the terrible literature search, Moenssens himself admits we were right:

> The *Exorcism* authors, in their self-directed, although partial and highly selective, literature study of handwriting identification had neglected to study the - of the worthy and unworthy handwriting analysts [sic]. They had stumbled upon some truly valid points that were made with professorial eloquence: (1) there was indeed a dearth of published empirical information relating to the proficiency of document examiners, as they had asserted . . . .

Moenssens, *Post*-Daubert *World, supra* note 1, at 300 (footnotes omitted). Elsewhere, Moenssens has been less grudging concerning the absence of empirical evidence pointed out by the *Exorcism* article: "On that the critics are absolutely correct." Andre Moenssens, Address to the AAFS, (Feb. 17-22, 1997) (speaking to an audience of QDEs, taken from Saks's contemporaneous notes). "Document examiners have not done the kind of empirical research that could have and should have been done." *Id.*

calculated error rates and statistics."[30] Further, he suggests that examination of the validity of practice in nearly any clinical field in ways that yield such data is well nigh impossible,[31] and therefore the assumption that such data ought to exist and be preferred to anecdotal declarations of self-belief is wrongheaded.[32] This position ignores the role of validity research in virtually every general field that is embedded in any enterprise that can credibly claim the mantle of science in any sense. As Moenssens points out, many practitioners undertaking clinical tasks in areas usually regarded as "scientific," such as surgery or psychology, may not, in much of what they do, be practicing something comfortably called "science."[33] However, as he fails to point out, in virtually any such field practitioners are supported within the general enterprise by a research arm that not only expands knowledge that might prove useful in the clinical setting (such as the development of a new drug), but also provides research on the validity and effectiveness of the techniques of practice. Moreover, the practitioners cooperate with the researchers to enable the research to be conducted. For instance, regular research is done on the effectiveness of various surgical techniques. As a result of this research, many such techniques have been shown to do more harm than good and have been abandoned.[34] There is no reason why similar research could not be done on handwriting identification, dowsing, or virtu-

---

[30] Moenssens, *Post-*Daubert *World*, *supra* note 1, at 279.

[31] Moenssens refers to "opinion testimony based on techniques that have not been *or cannot be* subjected to rigorous scientific methodology and validated by statistical information on error rates . . ." and also to "established forensic sciences that have not and *cannot be* measured and validated by mathematical and statistical models[.]" *Id.* at 282 (emphasis added). He further states that handwriting experts engage in judgments "*that defy quantification and measurement.*" *Id.* at 253 (emphasis added). Finally, and most startlingly, he says that often opinion evidence based on comparison of two items cannot qualify "as 'scientific knowledge,' since many examinations are of unique and singular events, *the accuracy of which cannot be readily catalogued as right-or-wrong, up-or-down, or measured and quantified statistically.*" *Id.* at 301 (emphasis added). By this he seems to have turned the issue of identification into Schrödinger's cat. But we are hardly dealing with quantum phenomena. First, if the "events" were truly "unique and singular" they could not be compared to determine identity of source, since they would share no similarity, and second, if anything demands to be treated as a clear case of "right or wrong" it would seem to be a conclusion concerning identity of source.

[32] *See id.* at 305.

[33] *See id.* at 277-78.

[34] One measure of a field's scientific vigor is its capacity to discover its errors and replace them with newer and better knowledge and techniques. *See generally* COSTS, RISKS, AND BENEFITS OF SURGERY (John P. Bunker, Benjamin A. Barnes, & Frederick Mosteller eds., 1977) (reviewing research findings on a variety of surgical techniques, finding that many popular surgeries actually were ineffective or harmful).

ally any area of claimed clinical skill or expertise.[35] Indeed, Moenssens himself celebrates one person attempting such research, Dr. Moshe Kam.[36] In any area asserting the right to the label "science,"

[35] Indeed, Moenssens is clearly aware of this when he wants to be. See his marshaling of the substantial empirical research showing the invalidity of the personality trait prediction aspects of graphology in *Post*-Daubert *World*, *supra* note 1, at 261 n.31. The number and range of subjective human decision-making tasks of the sort that Moenssens seems to argue cannot be studied empirically, yet which have in fact been studied empirically, is great and diverse. For further examples, see generally Robert D. Hoge & Theodore Coladarci, *Teacher-Based Judgments of Academic Achievement: A Review of the Literature*, 59 REV. OF EDUC. RES. 297 (1989) (reviewing the relationship of teachers' judgments of students' achievements to the students' actual performance); Irving M. Lane, Paul L. Damiano & Lorne M. Sulsky, *Determining Decision-Making Effectiveness Using NCAA Basketball Tournament Results*, 17 J. SPORTS BEHAV. 79 (1994) (explaining the comparative ability of six different methods of predicting the final rankings of college basketball teams); Robert Libby, Ken T. Trotman & Ian Zimmer, *Member Variation, Recognition of Expertise, and Group Performance*, 72 J. APPLIED PSYCH. 81 (1987) (studying how well loan officers could evaluate the financial profiles of firms and assessing the likelihood that they would experience bankruptcy); Douglas Mossman, *Assessing Predictions of Violence: Being Accurate about Accuracy*, 62 J. CONSULTING AND CLINICAL PSYCH. 783 (1994) (reporting a meta-analysis of 44 studies and 58 data sets of the accuracy of expert predictions of violence in mental hospital patients); Sonya M. Stevens, Douglas K. Richardson, James M. Gray, Donald A. Goldman, & Marie C. McCormick., *Estimating Neonatal Mortality Risk: An Analysis of Clinicians' Judgments*, 93 PEDIATRICS 945 (1994) (setting forth a study of the accuracy of doctors and nurses in assessing the likelihood of survival of seriously ill newborns); Daniel P. Sulmasy, Karen Haller & Peter B. Terry, *More Talk, Less Paper: Predicting the Accuracy of Substituted Judgments*, 96 AM. J. MED. 432 (1994) (assessing the ability of family members to discern the treatment preferences of patients). The specific task of observing complex and subtle stimulus patterns and making judgments about what one sees there — the very kind of cognitive task involved in comparing handwritings — is one of the most well-developed areas of research in other fields, with many important examples to be found in medical applications (such as evaluating the effectiveness of various diagnostic tests, such as CT scanning), *see* John A. Swets, *ROC Analysis Applied to the Evaluation of Medical Imaging Techniques*, 14 INVESTIGATIVE RADIOLOGY 109 (1979), military applications (such as evaluating the ability of people to read radar and sonar patterns), *see generally* W.P. Colquhoun, *Sonar Target Detection as a Decision Process*, 51 J. APPLIED PSYCH. 187 (1967); C.H. Baker, *Target Detection Performance with a Stationary Radar Sweep-Line*, 27 ACTA PSYCHOLOGICA 361 (1967), and basic research in psychology (such as studying human visual and auditory perception), *see generally* GEORGE A. GESCHEIDER, PSYCHOPHYSICS: METHOD, THEORY, AND APPLICATION (2d ed. 1985). Moreover, Moenssens sees no contradiction in saying such research cannot be done and then citing to it when it finds graphology wanting, or handwriting identification expertise to appear to exist. *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 261 n.31, 323-24 nn.312-16 and accompanying text.

[36] *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 313-16. "Celebrate" may be too weak a word. According to Moenssens, he is the "respected university affiliated research scientist Dr. Moshe Kam." *Id.* at 255 n.11. Kam, research scientist at Drexel University, "heads the data fusion laboratory" and has "written several theses involving probability and statistics." *Id.* at 312 n.276. He is "a true scientist," *id.* at 313 n.280, who acts "[w]ith the admirable restraint of a true scientist," *id.* at 313 n.279, who acts "in the proper 'scientific' manner," *id.* at 316, and conducts his stud-

one is not remiss in expecting to find such research literature. It is its non-existence that is surprising.

3. The claim that we erred in concentrating on quantified validity research and in excluding anecdotal case studies as data.

Moenssens implies that we were unaware of the mountains of published reports that support the accuracy of handwriting identification expertise when we published *Exorcism* in 1989.[37] On the contrary, as we noted then, our examination of the literature turned up a large number of such reports, which, in fact, we generally examined, but which we excluded from consideration in our article. As we said then, except for the reports we examined in the article, "[t]he vast majority of handwriting 'studies' reported in these publications consist of anecdotal observations, hypothesis creation, and speculation."[38] Indeed, in *Science and Nonscience* we discussed at some length the weaknesses of such reports of self belief based on anecdotal evidence.[39]

On one level, there is no doubt that anecdotal evidence derived from practice is "empirical" evidence of a sort. An issue in scientific evaluation of the claims of any field, from astrology to particle physics, is how much reliance to place on such evidence, and there is an answer which enjoys a fair consensus in the scientific community in general: "Not much."[40] Further, skepticism toward such evidence is

---

ies "according to the most rigorous scientific standards," *id.* at 324 n.316, etc.

[37] *See id.* at 301.

[38] Risinger et al., *Exorcism, supra* note 2, at 738 n.30.

[39] *See* Risinger & Saks, *Science and Nonscience, supra* note 2, at 41 n.100.

[40] The reason is that "a finding of fact is only as good as the methods used to find it." MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY § 2-1.1, at 48 (David L. Faigman, David H. Kaye, Michael J. Saks, & Joseph Sanders eds., 1997) [hereinafter MODERN SCIENTIFIC EVIDENCE.]. Normal scientists care intensely about methodology and research design. That is because the quality of the inferences that can be drawn from any empirical data is dependent upon the quality of the research design. For example, one step up from anecdotal evidence on the reliability scale are "case studies," yet these are regarded as having "such a total absence of control as to be of almost no scientific value." DONALD T. CAMPBELL & JULIAN C. STANLEY, EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS FOR RESEARCH 6 (1966). Put simply, scientists trust better designed studies more than they trust poorer studies (or no studies at all). Where research designs of varying degrees of quality (viz., from the weaker case studies or correlational studies to the stronger quasi-experiments and true randomized experiments) have been used to try to answer the same question, it has been found that the less well controlled the research, the more enthusiastic the conclusions, and vice versa. *See* John P. Gilbert, Bucknam McPeek, & Frederick Mosteller, *Statistics and Ethics in Surgery and Anesthesia*, 198 SCIENCE 684 (1977).

doubly justified in an area (such as handwriting identification and some other forensic disciplines lacking common non-courtroom applications) where normal practice does not provide unambiguous feedback concerning the accuracy of conclusions. This was the whole basis of our criticism, in *Science and Nonscience,* of Judge McKenna's analogy of handwriting identification to "harbor piloting" in his opinion in *United States v. Starzecpyzel.*[41] Unlike handwriting identification experts, harbor pilots have clear feedback as to whether they have been accurate in every case of practice — they either arrive safely at the right dock or they don't. In an area like handwriting identification, where such feedback is lacking,[42] it would seem that properly generated validity data would be not merely desirable, but necessary, and the exclusion of anecdotal reports similarly would be not merely justified, but fundamental.[43]

4. The claim that we misused or misinterpreted the data

Moenssens thinks we should have ignored certain studies or drawn from them other conclusions.[44] Once again, Moenssens re-

---

[41] 880 F. Supp. 1027 (S.D.N.Y. 1995); *see* Risinger & Saks, *Science and Nonscience, supra* note 2, at 33-34.

[42] *See* Risinger & Saks, *Science and Nonscience, supra* note 2, at 33. Consider also the following testimony in *Martin, supra,* note 16, by Robert J. Muehlberger, then chair of the AAFS Questioned Documents Section and vice president of the ASQDE, and one of the persons credited by Moenssens with helping with *Post*-Daubert *World*:

> Q.In a Laboratory setting where you have a real life problem, how do you decide that your work is valid?
> A. It's by a review process. Another examiner would examine that case, that problem, basically utilize the reasoning and the methodology we utilize in that laboratory. And if they reach the same conclusion as the first examiner, then we would consider that in fact that those conclusions are valid and reliable.
> Q. And there is no other way to tell whether or not you are correct?
> A. In that situation, no.

Id. at 127-28.

> Q. Okay. Now, how many times do you think you have been wrong when you have made a comparison between a questioned document and a known sample?
> A. I wouldn't know.
> Q. You don't have any idea?
> A. Very few times, I would hope.
> Q But you don't know?
> A. Exactly, no.

Id. at 141-42.

[43] Anecdotal reports function at best as "investigative leads," to borrow a phrase from criminal investigation, which are weak evidence in themselves but may properly be used to focus a search for more dependable evidence in a particular direction. At best, they frame questions without providing answers.

[44] *See* Moenssens, *Post*-Daubert *World, supra* note 1, at 300-01. Specifically,

peats criticisms contained in the Galbraith article,[45] criticisms we dealt with in detail in *Science and Nonscience*.[46] Everyone concedes that the 1939 Inbau study was so flawed that it provided no meaningful data on experts' abilities, or their marginal advantage over lay persons, which was our original conclusion.[47] Moenssens criticizes us for including it at all, since Inbau claimed in an unpublished post-*Exorcism* memorandum that the 1939 study was not designed to generate such data.[48] Nevertheless, Inbau did not say this in 1939. Test results by experts were generated and included in the study as published, and we would have been properly criticized for *failing* to examine it.[49] As to the FSF studies, we faced two options in dealing with them: to dismiss them as meaningless (which would have resulted in a conclusion that there were no data whatsoever bearing on the validity of handwriting expertise, and also in the proper criticism that we had failed to present and examine all the available potential data), or to examine them as potentially meaningful. After all, while it is true that the FSF said that it regarded the data as inappropriate from which to draw any conclusions, it seems unlikely that it could have

---

Moenssens points to Fred E. Inbau, *Lay Witness Identification of Handwriting*, 34 ILL. L. REV. 433 (1939) [hereinafter "The Inbau study"] and the studies we have referred to collectively as the FSF studies. *See* Moenssens, *Post-Daubert* World, *supra* note 1, at 300-01.

[45] *See* Galbraith et al., *supra* note 25, at 9-11.

[46] *See* Risinger & Saks, *Science and Nonscience*, *supra* note 2, 49-53. Moenssens adopts a curious tactic in dealing with *Science and Nonscience*. Regarding that article, in footnote 7 of *Post*-Daubert *World*, he says, "No attempt will be made herein to make a detailed analysis of the *Science and Nonscience* article since Dr. Saks graciously provided this author with a copy of it only a week prior to finalizing the copy of this article." Moenssens, *Post*-Daubert *World*, *supra* note 1, at n.7. Moenssens then goes on to criticize *Science and Nonscience* selectively at numerous points, while ignoring other points made in the article that are apposite to his claims. For the record, Moenssens was sent a copy of *Science and Nonscience* in late May or early June of 1997, *one year* before his article was published.

[47] *See* Risinger et al., *Exorcism*, *supra* note 2, at 741. The "everyone" includes Kam, *see* Kam I, *supra* note 8, at 6, the Galbraiths, *see* Galbraith et al., *supra* note 25, at 9, and Moenssens himself, *see* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 300.

[48] *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 301 n.222.

[49] For some reason, it appears important to Moenssens to argue that the Inbau study was not a "validity study." He wrote:

> The *Exorcism*... authors begin their 'destruction' of handwriting identification lore by discussing the so-called 'Inbau Study' of 1939, which they called a 'primitive and flawed validity study from nearly 50 years ago.' The gross mischaracterization of this study alone as a validity test of the abilities of document examiners casts serious doubt on the research methodology of the *Exorcism* authors.

Id. at 301 n.222 (quoting Risinger et al., *Exorcism*, *supra* note 2, at 738). Moenssens is simply in error in his characterization. Flawed though it is, the Inbau study meets all the criteria of a "validity study." *See infra* at notes 245-50 and accompanying text.

regarded the exercise as totally without meaning, or it would have ceased to give the tests or refrained from presenting the data.[50]

There are always various ways of dealing with raw data once generated, and there are always various ways of aggregating data from separate studies. In both *Exorcism* and *Science and Nonscience*, we considered various aggregation strategies, made elections, and spelled out the assumptions behind those elections.[51] The Galbraith article attacked some of those strategies, to which we responded in detail in *Science and Nonscience*.[52] One striking thing about the Moenssens article is that, unlike the Galbraith piece, it does not attempt to make *any* substantive criticism of our data analysis, beyond characterizing it conclusorily as "'inappropriately-cited percentages and averages, surrounded by deviations and variances, and garnished with a dollop of Bayesian theory' that provides the veil of legitimacy for otherwise highly questionable research," quoting a sentence from the Galbraith article.[53] What Moenssens fails to reveal is that the Galbraiths generated their own data analysis table on the same FSF data dealt with in *Exorcism* and found that the experts failed to exceed *chance* (that is, sheer guessing) on two of six subtasks in those tests.[54] Suffice it to say

---

[50] This point was made in *Science and Nonscience*, *supra* note 2, at 42. See the observations on possible political motivation, *infra* note 257 and accompanying text.

[51] *See* Risinger et al., *Exorcism*, *supra* note 2, at 747-50 (and accompanying footnotes); Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 43-53 (and accompanying footnotes).

[52] *See* Risinger & Saks, *Science and Nonscience*, *supra* note 2 at 49-53. The main substantive disagreement between the Galbraiths and ourselves involves how to treat reports labeled "inconclusive" both by the test takers and the FSF, which were sometimes accompanied by narrative statements leaning toward one writer or another. Our view of the matter parallels the one arrived at by the court in *United States v. Starzecpyzel*, 880 F. Supp. 1027 (S.D.N.Y. 1995), after this very issue was considered in the *Daubert* hearing in that case:

> The document examiners reached the correct answer 52 percent of the time. The lay persons reached a correct answer 50 percent of the time. The Galbraiths then took the document examiners' answers and the lay people's 'inconclusives,' and they fiddled with them . . . . And then came up with a more impressive difference.

Id. at 1037.

[53] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 306 (quoting Galbraith et al., *supra* note 25, at 15). The uncareful reader might think this comment was a criticism directed at something that actually existed in the *Exorcism* article, but it was instead a vague and unaimed shot fired by the Galbraiths at nothing in particular. The Galbraith article did not actually apply the quoted description of what they called "statistical innumeracy" directly to us, and rightly so, since, for instance, we didn't garnish anything in either article with a "dollop of Bayesian theory." Moenssens has reloaded the same missile and fired it again in our general direction without hitting anything.

[54] Galbraith et al., *supra* note 25, at 15 tbl.3, reproduced in *Science and Nonscience*, *supra* note 2, at 52. Moenssens did not think this remarkable finding to be worthy of

that we believe the fair-minded reader who considers what we actually wrote and how we approached the data will conclude that we were quite justified in our conclusions in 1989: that, while future tests might "show document examiners faring better," in light of the FSF studies the corpus of available evidence on the existence of handwriting identification expertise went "from no data to negative data" and that in its most charitable light and taking into account the flaws in the FSF studies, the most that could be said at that time was that "no available evidence demonstrates the existence of handwriting identification expertise."[55]

One final observation is in order in this section. Throughout the Moenssens article there is the clear implication that if we had limited our 1989 observations to non-Osbornian handwriting experts we would have been justified, but the real sin of *Exorcism* was our inclusion of Osbornians in our global conclusions.[56] This position overlooks the fact, implied in *Exorcism* and made explicit in *Science and Nonscience*, that all the available data, meaningful or not, have been generated by tests run on Osbornians.[57] Therefore there are still no data explicitly on graphologically oriented practitioners. Testing might show them doing worse, but it remains a possibility that it might show them doing better. Such a result might be counterintuitive, because of the claimed superiority of Osbornians in training, experience, and method, but there are counterintuitive precedents for such things in other fields and some counterintuitive data in this field.[58] We remain firmly agnostic.[59]

---

sharing with his readers, though the *Starzecpyzel* court took note of it. *See Starzecpyzel*, 880 F. Supp. at 1037.

[55] Risinger et al., *Exorcism*, *supra* note 2, at 750-51.

[56] *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 264 n.41, 303 n.231 & 307 n.248.

[57] To date, all studies, including the AAFS studies and both Kam studies, have been run on document examiners from government laboratories, and "since government laboratories tend to be officially Osbornian, if there are few data on the dependability of identification of authorship by Osbornian practitioners, there are none on the dependability of graphologically oriented practitioners." Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 298 n.210. Moenssens apparently agrees about the population of the tested universe:

> A sizable number of the traditional questioned document examiners are employed in crime laboratories and in local, state or federal law enforcement units. However enough of them are retired from public service and have entered into the private practice of forensic document examination so that defense lawyers have equal access to their services. By contrast, most of the graphoanalysts are in the private practice, either full time or part time.

Moenssens, *Post*-Daubert *World*, *supra* note 1, at 264-65.

[58] There is reason to believe that the accuracy of predictions of future danger-

B. *The Kam Studies*

Professor Moenssens entertains not an iota of doubt concerning handwriting identification dependability based on the data generated by the FSF tests reported in *Exorcism* (or the further negative data from 1988 and 1989 FSF tests reviewed in *Science and Nonscience*); nor does he entertain any doubts that experienced Osbornian experts are substantially more accurate in their identifications than virtually any untrained juror. However, Professor Moenssens takes the position that, were any such doubts ever justified, they have been totally and completely dispelled as the result of the work of Dr. Moshe Kam, specifically his 1997 study (Kam II).[60] Moenssens asserts that this study "laid to rest, for all objective reviewers, the debate over whether professional document examiners can reliably identify authors of handwriting and possess a skill that is absent in the general population,"[61] and "shows conclusively that questioned document examiners possess a skill in identifying writers of documents that is not possessed by lay persons."[62]

In this, Professor Moenssens is triply misled: First, no single study can show anything that conclusively;[63] second, the Kam II study

ousness is either unaffected or negatively affected by training in psychiatry. See the authorities collected in *Exorcism*, *supra* note 2, at 750 n.80. *See also* John Monahan, *Clinical and Actuarial Predictions of Dangerousness*, *in* MODERN SCIENTIFIC EVIDENCE, *supra* note 40, § 7. And in regard to accuracy of handwriting identification, the FSF data showed no significant correlation between accuracy and time on task, and only marginally significant correlation between accuracy and years of experience. *See* Risinger et al., *Exorcism*, *supra* note 2, at 785-87 app.2.

[59] Note that we see no reason to believe that graphologically oriented practitioners are any better, and would certainly not be surprised if they were worse. Nevertheless, observations similar to the ones above have prompted Moenssens to imply that somehow we are supporters of the graphologists. It is at that point where we begin to feel we are being attacked more for heresy than anything else.

[60] *See* Kam II, *supra* note 8.

[61] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 314.

[62] *Id.* at 324. The cited quotations are but two of a series of encomiums paid Dr. Kam by Moenssens. *See supra* note 36.

[63] A well-established understanding among empirical researchers is that replication and extension are necessary to the development of knowledge, first to establish the generality of a finding (can it be repeated? does it hold beyond the narrow set of circumstances of any one study?) and, second, to find the boundary conditions beyond which the finding does not hold. To expect any one study to accomplish so much is extravagant. *See* ROBERT ROSENTHAL & RALPH L. ROSNOW, ESSENTIALS OF BEHAVIORAL RESEARCH 111, 182 (1984). According to Rosenthal and Rosnow, "Replication is science's empirical system of checks and balances to expose and reduce errors of observation," and "The crucial role of replication is well established in science generally . . . ." *Id.* On many, if not most, important questions, dozens, if not hundreds, of studies can be found exploring various aspects of these questions. It should be noted that the United States Department of Justice apparently does not share Moenssens' view that things have been laid totally to rest. In June of 1998 the

had several serious flaws, which leave open questions as to its actual meaning;[64] and third, even taken at face value, the study does not mean what Moenssens (or Kam, for that matter)[65] seems to claim for it.

No matter how hard the forensic science community wishes for a single magic bullet, which will dispatch all doubts, science just does not work that way, especially in an area like handwriting identification where actual practice involves a range of tasks widely varying in difficulty. A single, unreplicated study does not establish much under any circumstances,[66] and no single test can map the terrain of dependability in a complex area such as handwriting identification.[67] This is not to say that such individual studies are valueless — far from it. These studies mark the beginnings of real knowledge. While we have been searching in our criticism of both the Galbraith study

---

National Institute of Justice (the Justice Department's research agency) issued a solicitation for research proposals on "Forensic Document Examination Validation Studies." The solicitation document says that the procedures of document examiners "must be tested statistically in order to demonstrate that following the stated procedures allows analysts to produce correct results with acceptable error rates." NATIONAL INSTITUTE OF JUSTICE, SOLICITATION, FORENSIC DOCUMENT EXAMINATION VALIDATION STUDIES (June 1998).

[64] Kam in fact is engaged in a *program* of research, with each study trying to improve upon and fill in the gaps of the previous one. Criticisms of Kam I led to the improved research design of Kam II. Criticisms by Denbeaux and direct communications between Saks and Kam about the design flaws of Kam II led to the tests that were conducted for Kam III. *See* Deposition of Moshe Kam at 48-49, United States v. Gilreath, No. 96-CR-472 JTC (N.D. Ga. Dec. 4, 1997) (unpublished, copies on file with authors) [hereinafter *Gilreath* depositions]. This is an example of how knowledge development in science proceeds (notwithstanding Kam's extravagant pronouncement in Kam II of the end of history in this area of inquiry. *See infra* note 65.).

[65] In Kam II, Dr. Kam makes the following statement, which is an extraordinarily extravagant statement to be made by a serious scientist reporting the results of a single study: "[T]he results of our test lay to rest the debate over whether or not professional document examiners possess writer-identification skills absent in the general population. They do." Kam II, *supra* note 8, at 778 and repeated in substance on page 785. This is the same article that Moenssens has characterized as showing "the admirable restraint of a true scientist." Moenssens, *Post*-Daubert *World*, *supra* note 1, at 313 n.279. *See also supra* note 36.

[66] *See supra* note 63.

[67] Kam himself is well aware of this. When asked by Judge Carnes in the *Martin* case if evaluating individual document examiner competence would require "a fuller battery of tests" than the test performed in Kam II, Kam replied:

> Yes. If I were requested not to answer the questions that I have answered but to try to answer the question what is an appropriate set of tests in order to tell that someone is good enough at a certain reliability level, I'll have to work on it. I mean, this doesn't answer it. I'll have to work on it. It would be a study bigger than this.

Martin Transcript, *supra* note 16, at 237.

(which Moenssens rather unaccountably fails to describe) and the first Kam study (which Moenssens also fails to describe) in our previous writings,[68] we have, we believe, been appropriately open in stating what these two studies taken together tend to suggest tentatively, at least as to the very limited and artificial tasks tested: the possible aggregate superiority of experts as a group in avoiding false positive attributions under test conditions,[69] and the heavily bimodal performance of non-experts, with the top half of non-experts approaching the bottom portion of the experts in performance even as to false positives.[70]

Which brings us to Kam II, and why it cannot deliver what Moenssens claims for it. Between the beginning of May and the end of September 1996, Kam, Fielding, and Conn administered a test of their own design to three groups of about thirty-five questioned document examiners each, a group of eight document examiner trainees, and a group of forty-one untrained non-experts.[71]

The test materials were generated as follows: 150 persons, ages twenty to twenty-seven, were selected by an unknown protocol to provide writing samples, and they agreed to do so. Each writer worked on a wide and well-lit table in a classroom setting. Each writer generated twelve documents on eight and one-half by eleven inch, twenty pound white paper, copying three short assigned texts four times each. Each writer used pens supplied by the test designers, both blue medium Bic pens and black medium Bic pens, and each was told to use both colors, switching "every 2-3 documents," so each writer created both blue and black documents in no exact fixed ratio or order.

---

[68] *See* Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 58 (discussing Galbraith study), 59-63 (discussing Kam I). In attacking this evaluation of Kam I, Moenssens typically characterizes it as a "denigration" "by a non-scientist lawyer of the work of a true scientist." Moenssens, *Post*-Daubert *World*, *supra* note 1, at 313 n.280. Saks, who stands behind our discussion of Kam I, takes umbrage at this characterization — he has always regarded himself as a non-lawyer scientist. For a discussion of Dr. Saks's qualifications, see *infra* note 240.

[69] It cannot be stated too often that any superiority on the part of document examiners under test conditions, if it truly exists, will not necessarily carry over to actual practice, where the pressures on document examiners, as on other forensic practitioners, are to make matches that confirm positions already arrived at by other investigators, a fact recognized by Moenssens himself in other writings. See the quotations collected *infra* at note 283.

[70] *See* Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 61. No one could begin to guess this evenhandedness of our evaluation from reading Moenssens' article.

[71] *See* Kam II, *supra* note 8, at 779-81. All details of Kam II test procedures are drawn from pages 779-81 of that eight-page article. Because the material is so short, specific footnote references to pages are omitted as unnecessary and burdensome to the reader.

This resulted in 1800 documents, four each of three texts by 150 writers. Thirty of those writers were then selected (presumably at random) and all the documents generated by those thirty (360 documents) were placed in a set. These documents were then coded with random numbers for writer identity. Random documents were then drawn from the set of 360 until six documents by six different writers were obtained (that is, after drawing the first document, if the second document was by the same writer, it was returned to the pool and another was drawn, and so forth, until a set of 6 documents by 6 different writers was obtained). These documents could represent randomly any text and either color ink. This set of six was then labeled "Unknown A1." The same process was repeated until twelve such sets were obtained (Unknowns A1 through A12) which together contained seventy-two documents in twelve sets of six each, each of the six by a different writer in each set. The remaining 288 documents were then randomly distributed into twelve sets referred to as "database packages," (database A1 through A12), each containing twenty-four random documents.[72] A similar process was undertaken with the remaining documents not in set A, generating a set B, with the same subset characteristics, and so on until there were five such universes, A-E, containing twelve six-document "Unknown" sets each and twelve twenty-four-document "database" sets each. Each test participant was tested by randomly selecting a Universe A-E, then within that Universe randomly selecting an "unknown" set from the twelve, and randomly selecting a "database" set from the twelve. The researchers then explained to the test subject that the "unknown" sets contained six writings by six different writers without disguise, and asked the test participant to determine whether any document or documents in the "database" set of twenty-four were written by any writer represented in the "unknown" set.

The report of the study is a bit hazy in setting forth the limiting conditions of the sets so generated, therefore, we have a problem in figuring out the specifics on this point. It is inexplicit whether, once

[72] The use of the label "unknown" was somewhat problematic. Among document examiners, "unknown" is the label given to a document of unknown authorship to be compared with standards of known authorship. In the Kam II test, none of the documents in either "unknown" sets or "database" sets were unknowns in this sense. They were in fact functionally the same from that perspective: Both sets were known to the testers and unknown to the test takers. The test was thus the same kind of sorting test involved in Kam I. Moenssens seems to have been misled to some extent by the label, because he refers to the documents in the database sets as "knowns," when in fact they were no more "knowns" than the "unknowns." *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 314

"Unknown A1" was generated, the writers there represented were removed, so that set A2 would have six writers who were different from each other and also all different from the ones in set A1, or whether the process was simply repeated on all the *documents* remaining after the generation of set A1, in which case set A2 could, in theory, contain the exact same writers as A1. It appears that the latter must be the case, because with twelve sets of four, one would need forty-eight writers to have no overlaps, and there were only thirty writers in the pool. In the latter case, it is possible (though unlikely) that each set A1 thru A12 represented exactly the same six writers, since each writer had generated twelve documents. Similarly, in regard to the "database" sets, in the event of the distribution just described, *a fortiori*, no "database" set would, or could, contain a document truly matching one in an "unknown" set. This theoretical (and very remote) possibility is only important to understand the real meaning of the first three tables in the Kam II report. These tables are obviously descriptive of the distribution of matches in the actual tests as given to the participants, not of the probability of such matches resulting from the distribution process described above. Thus, while there is a statistical probability that some test might be administered that had no true matches, in the 154 tests actually run there was always at least one true positive match. Similarly, at the other extreme, though there is a remote possibility of as high as eighteen true matches, in the actual 154 tests administered there were never more than ten.

Before we look at the published results of these tests, let us examine the test design, and the limits of what we might expect that it can and cannot tell us. First, like Kam I, this is a multi-document sorting test of a type encountered rarely, if at all, in actual practice.[73] This is not to say the results are meaningless. As we said in *Science and Nonscience* in regard to the similar characteristics of Kam I, "the ability to accurately perceive diagnostic patterns of similarity and difference in the writing represented by the test materials would likely be common to both the test and to many kinds of real-life problems."[74] The limitations of the Kam II test are mentioned merely to

[73] Kam states that the "format of the test was selected to resemble a multi-suspect case in which extensive examination of documents was required," but in fact, because there were multiple documents in each comparison set, none of which were known as to origin by the examiner, it resembles no real case. Kam II, *supra* note 8, at 780. Perhaps it would be better if real cases were presented to examiners in this blind a fashion, but they are not. *See* Risinger & Saks, *Science and Nonscience, supra* note 2, at 64.

[74] Risinger & Saks, *Science and Nonscience, supra* note 2, at 60.

caution against simple extrapolation to actual practice without further thought.

Second, there are problems of motivation with regard to some study participants. Professional document examiners would be expected to put in more focused effort, given that their careers (and the fate of their profession) are in some sense on the line, than would graduate students merely participating in a study. This gap in the motivation of study participants was a serious problem in Kam I.[75] However, in Kam II, the study designers have made a commendable attempt to correct for this factor by offering the non-experts a schedule of monetary rewards and penalties. Unhappily, the schedule of rewards and penalties utilized was particularly unfortunate and problematical. The non-experts were told that they would receive a twenty-five dollar participation fee, and that their payment would not go below this regardless of actual performance. They were then told they would receive an additional twenty-five dollars for each "true positive" match, that they would lose twenty-five dollars for each "false positive" match,[76] that they would lose ten dollars for failing to see a true match, and that they would gain nothing for accurately rejecting a non-match. This payoff schedule may be schematically represented by the following matrix:

[75] *See id.* at 61.

[76] Note that under the stated reward scheme, if you had no idea of whether any of the documents in fact matched, the incentive would be to take one random guess of a match, which would be free since you couldn't lose the $25 participation fee, and you might get lucky.

		Reality: Match	Reality: Non-Match
Decision:	Match	True Positive +25 (Very Good)	False Positive -25 (Very Bad)
	Non-Match	False Negative -10 (Bad)	True Negative 0 (Indifferent)

Under this regime, if the non-expert participants' objective was to maximize their payoff, they would guess a match whenever it appeared to them to be as likely as a non-match. This is because, over the long run, it would appear that such a strategy would at least break even, whereas guessing "no match" on such an evaluation would lose an average of five dollars per guess over the long run.[77] Plus, there was no real incentive to avoid false positives for fear of actually losing something they already had at the beginning of the test because they were guaranteed the twenty-five dollar participation fee no matter how bad their performance. This reward schedule seems guaranteed to make the non-experts risk-preferring regarding finding positive matches, even in the face of instructions that they should declare a match only if they were absolutely sure. The effect of the payoff structure compounds the well-known tendency for a majority of people to become risk-preferring in circumstances of potential high rewards and low costs, regardless of rational odds (sometimes referred to as the lottery effect).[78]

[77] This strategy would be expected to lose $10 half the time (when the guess was wrong) and gain nothing the other half (when the guess was right).

[78] *See* Daniel Kahneman & Amos Tversky, *Prospect Theory: Analysis of Decision Under Risk*, 47 ECONOMETRICA 263 (1979); Lola L. Lopes, *Remodeling Risk Aversion*, *in* ACTING UNDER UNCERTAINTY: MULTIDISCIPLINARY CONCEPTIONS 267 (George M. von Furstenberg ed., 1990); Lola L. Lopes, *When Time Is of the Essence: Averaging, Aspiration, and the Short Run*, 65 ORG. BEHAVIOR & HUMAN DECISION PROCESSES 179 (1996); Amos Tversky & Daniel Kahneman, *Advances in Prospect Theory: Cumulative Representation of Uncertainty*, 5 J. RISK & UNCERTAINTY 297 (1992).

In contrast, the document examiners entered the test under quite a different effective payoff schedule. Foremost, they knew that the worst thing for a professional document examiner to do on any proficiency test is to commit a false positive error.[79] (This risk averseness to false positives on known tests is not necessarily present in normal practice, for a variety of reasons discussed *supra*, at note 69). Secondly, for the document examiners, a correct detection of a non-match is also a highly desirable indicator of affirmative expertise. The incentive matrix for document examiners taking the Kam II test might be represented as follows:

		Reality: Match	Reality: Non-Match
Decision:	Match	True Positive (Very Good)	False Positive (Extremely Bad)
	Non-Match	False Negative (Bad)	True Negative (Good)

Thus, in a situation of equipoise, this reward structure would impel experts toward declaring a non-match, and the non-experts toward declaring a match.

Hence, the experts and non-experts took the tests under different incentive structures, which would be expected to yield more false positives for the non-experts[80] even under equally accurate probability judgments about authorship.[81]

---

[79] *See* Moenssens, *Post-*Daubert *World, supra* note 1, at 315. Such a false positive in the real world can result in the conviction of an innocent criminal defendant, a result the official ideology of our criminal justice system disvalues much more than an inaccurate acquittal, as reflected in the requirement of proof beyond a reasonable doubt. Thus, for an expert to commit such an error on a proficiency test undermines the status of the expertise in the eyes of the law, and therefore must be avoided at all costs *in tests*. Whether a false positive must be as stringently avoided in practice is a different issue. *See supra* note 69.

[80] To correct for this, future tests should impose high disincentives on the non-experts for false positive responses, substantially higher than the reward for true

Third, and somewhat unaccountably, is the "ink color" variable, also present in Kam I. It is not clear why the writers of the exemplars were asked to use two different color inks. The effects of the different inks on the results, or on the aggregate results between groups, cannot be assumed *a priori* to be trivial. As we said in *Science and Nonscience*, this "introduced a variable into the study's design having unknown impact and no apparent relevance to the issue of identification from form."[82] Nevertheless, this additional variable is probably

---

positives, and equalize the gain and loss for true negatives and false negatives (missed matches).

Scientists who study decision-making of the sort engaged in by forensic scientists, an area of research called "Signal Detection Theory," have developed ways of measuring the raw acuity of sonar operators, radiologists, or document examiners separate from the subjective threshold that an observation must cross in order for examiners to decide that what they have observed is a submarine or a tumor or a handwriting "match." Such research has found that incentives for preferring to err in one way (a false negative: failing to detect a tumor) rather than another (a false positive: seeing a tumor where there is none) typically have considerable impact on where the psychological threshold is placed, regardless of the raw perceptual accuracy of the examiner. *See* John A. Swets, *The Science of Choosing the Right Decision Threshold in High-Stakes Diagnostics*, 47 AM. PSYCHOLOGIST 522 (1992). (Note that because of its human component, this research is significantly different from its related computerized cousin, Pattern Recognition, in which Dr. Kam is an expert. *See* Moenssens, *Post-*Daubert *World*, *supra* note 1, at 312 n.276.)

[81] In a recently published Article, *The Effects of Monetary Incentives on Document Examination Professionals*, 43 J. FORENSIC SCI. 1000 (Sept. 1998) [hereinafter Kam III], Kam et al. administered the Kam II test to four sets of laypersons under four different incentive schemes and found no significant difference in result among the four test groups. We make the following observations:

1. None of the incentive schemes tested had the characteristics outlined in note 80, *supra*.

2. The findings are surprisingly at odds with the general results of similar incentive studies in other areas. *See supra* note 80.

3. While there appeared to be no significant differences among the four newly tested lay groups attributable to the different incentive schemes, the aggregate performance of the newly tested lay groups showed a 41% improvement in avoiding false positives compared to the original Kam II lay group. (The wrong association rate dropped from .383 to .227.) This is a startling volatility of result, given the same task, and it raises substantial questions about the reliability of all of Kam's lay person test results. Kam tries to argue that this precipitous (and statistically significant, $Z = 1.85$) decrease in false positives is accounted for by a general tendency of Kam III subjects to declare fewer matches of any kind, giving them "a lower (worse) hit rate." Kam III, *supra*, at 1000. However, Kam fails to perform any significance tests to back up this claim. In fact, there is no significant difference in hit rate between Kam II and Kam III non-professionals ($Z = 0.034$, p is nonsignificant).

Professor Denbeaux has noted this volatility of result in testimony concerning a draft of Kam III, but this testimony was not included in Kam III's rather selective Appendix of Denbeaux's testimony.

[82] Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 59 n.142.

best regarded as simply another subtask variable of the type inevitably present in the tests as designed, for reasons explained below.[83]

Fourth, a much more serious problem with the Kam II study is the possibility that some of the document examiners (but not the non-expert participants) had helpful information about the test in advance of its administration.[84] Apparently, Kam allowed the earlier document examiner test subjects to review the results of their tests at some point before the next document examiner groups were tested.[85] Now, Kam claims to have protected against this having any effect by making sure that no pairing of a set of "unknowns" and a set of "database" documents was ever used twice. However, it should be obvious that quite a lot of useful test design information might be gleaned in such a session. For instance, if you know how the data base was generated, you can figure out pretty quickly how poor in true matches the tests are likely to be, that is, what the rough probabilities of maximum and minimum numbers of true positives are. Such information was not available to the non-experts, who might be expected to assume a universe much richer in matches, which would encourage guessing. Furthermore, while the document examiners were divided into three regional groups for statistical purposes, the members of at least the Northeast group were tested in at least two different sub-groups at two different times, with their results being aggregated statistically. Thus, there were four different physical groups of experts tested at four different times, and the first group might have had valuable information that could be passed to the members of the other groups. It seems especially unnecessary to

---

[83] *See infra* text accompanying and preceding note 86.

[84] *See* Risinger et al., *Exorcism, supra* note 2, at 62. (discussing this problem in regard to Kam I, which in that study rendered it impossible to know if individual performances were in fact group performances in regard to the document examiners tested).

[85] It is not fully explicit that this occurred. Kam says that, because "unknown" and "database" pairings were never repeated,

> Even if correct results from an early test were fully known to all test-takers in a later session, this information was practically useless. We do not believe that any attempt was made to record or share results from our tests between test-takers. However, if such attempts were made, they could not affect the results in a meaningful way.

Kam II, *supra* note 8, at 779. We take this to mean that results were provided the test takers at some point in advance of the next group of tests. In order to make sure this occurred, Saks has repeatedly requested explicit clarification from Kam, by e-mail, telephone, and by regular mail, but Kam has not responded. Under these circumstances, we have proceeded on the assumption that, based upon Kam's statement, at some point test subjects reviewed the results of their tests. This is one more illustration of the hidden pitfalls of human testing that may not be immediately apparent to the researcher more accustomed to computer modeling.

have run the tests in such a way as to create this problem, since all that was necessary was to collect all the data before conducting any post-mortem analyses.

Fifth, and finally, the actual tests presented an unknown variety of subtasks in an unknown distribution. In addition, some important varieties of subtask were clearly absent from the test, and as to those, the test can generate no direct data of any kind. Let us expand a little on this. Like the Kam I universe of writing exemplars, the much larger and more controlled universe generated for Kam II was generated with no attempt at isolating subtasks (for instance comparing two exemplars with similar but unusual "class characteristics," such as the handwriting of two German immigrants of similar age and sex). Any such subtask *might* have been present somewhere in some run of the test, but we do not know which were, which were not, or when. This is not necessarily a criticism, but it must be kept in mind when determining what the test can tell us. Each test was a little different, presenting a different set of challenges, and the most that can be said is that the challenges are likely to be typical of a certain range of comparison situations. They are also unlikely to mirror those encountered in litigation. This is because many of the subtask problems are likely to be trivial, like distinguishing between two randomly selected humans of all races and sexes, since the exemplars were generated with no apparent attempt to make the set of materials richer in confusingly similar exemplars than random life would be.[86] Most importantly, some very critical subtasks were not under test at all, such as the effects of forgery or disguise, the particular problems of signatures or adolescent handwriting, or those of the elderly or infirm, or the like. However, with these limitations in mind, and within the range of aggregate subtasks present in the tests, the tests can generate, and would be expected to generate the following important types of data:

1. The aggregate performance of the professional document examiner group as to true positives, false positives, true negatives, and false negatives.
2. The aggregate performance of each subgroup of professional document examiners as to the same categories of result.
3. The aggregate performance of the trainee group as to the same categories of result.

---

[86] *See* Risinger et al., *Exorcism, supra* note 2, at 64 (for a discussion of the same problem in Kam I).

4. The aggregate performance of the non-expert group as to the same categories of result.
5. The distribution of performances among the professional document examiner group (best score, worst score, distribution in between, as to all categories of result).
6. Same for sub-groups.
7. Same for trainees.
8. Same for non-experts.

Indeed, these data are so important that, with a small universe of 154 test administrations, one might expect to see a frequency distribution, or even a data table giving each taker's (unidentified by name) individual scores, as was done in Kam I. In the Kam II report, however, all that is given is the aggregate performance of groups, and that only in numerical averages. There is no distribution information given at all, even in the form of standard deviation values. It's not that this information does not exist or was not available to Kam. It just is not given in the report.

Think of what this means. We do not know how well the best lay people performed, or how poorly the worst document examiners performed. Kam knows, we don't. The aggregate data establish that there is no significant difference in the average performance of experts and non-experts in regard to true positives.[87] It is only in regard to false positives, admittedly the more dangerous type of error, that document examiners have a significant aggregate advantage. However, without distribution information, we do not know if, for instance, the better half of the non-experts is as good or better at avoiding false positives as the worse half of the document examiners.

The above example is not as farfetched as it might sound. The data from the Galbraith study and from Kam I tended to indicate that non-experts were bimodal in their accuracy distribution, including false positives. If that trend were to hold in Kam II, the best of the non-experts could still be better than the worst of the experts, but the average for non-experts would be dragged down by the really poor performance of the worst of the non-experts. Similarly, the high average for document examiners could conceal a cluster of really poor performers on the low end.[88] These data exist but were not pub-

---

[87] Note that the proper conclusion is "no significant difference," not "accept" the null hypothesis as Kam mistakenly concludes in Kam II. *See* Kam II, *supra* note 8, at 784 tbls. 10, 11, 13, 15, and accompanying text. A null hypothesis is only a starting point for a scientific study. Based on evidence, it can be rejected. But the failure to reject it as false is not the equivalent of "accepting" it as true.

[88] We are well aware that the test design, which resulted in each test adminis-

lished. What is worse, Kam has repeatedly refused requests from us and others to produce these data.[89] Under such circumstances

tered being different from every other, and each test participant therefore taking a somewhat different test, makes statements about comparative individual performance problematic, absent some way to rate the relative difficulty of particular tests. Nevertheless, other moments of the distributions have as much claim to meaning from these data as the one Kam chose to publish, but these cannot be derived without the data.

[89] Kam originally agreed to share his data in January of 1997:

> Q. In your 1994 tables [from Kam I], you actually could tell if some students made a lot of mistakes versus very few?
> A. Yes.
> Q. Okay. Can you tell that from the new studies?
> A. From the new, no. But because of the fact that I imagine that this will be of interest to many people, I will make it available. The only problem is that if I send to the journal a table of 147 entries, I mean I don't even know how to print it. I'll make — I mean I'll make it available.

Martin Transcript, *supra* note 16, at 217. Kam then goes on to caution against unsophisticated comparison of individual performances from the raw data because, as we have noted *supra* note 85, each test was a little different, presumably some much harder and some much easier. We, of course, intend no unsophisticated comparisons.

No data were ever forthcoming in response to requests, however, and by March of 1998, Kam explicitly changed his position and now refuses to share his data:

> Q. Do you have other concerns about, and again we've heard the complaints, essentially, that your data is not being shared. Do you have other responses to that criticism?
> A. I keep hearing that I should share my raw data and I would like to know, first of all, what is meant by that? If it means that I have to go and show these, I object to that. I have a proprietor recording on these things. I mean, there is tests designed — I mean unless — I mean if the Journal of Forensic Science want to order me, that's absolutely fine. But to go just because someone wants to look at these things because they have some hypothesis of some need to show that what I have done is wrong and go ahead and distribute to the whole world how I do my testing and how I do the coding, I object to that. Secondly, I'm not done. There are some things I want to know from this data that I either couldn't finish yet or I don't have right now the funds to do and I am soliciting funds to do. Let me give you one example. I want to know, in this test and some cases, the examiner had the same text to compare and sometimes the examiner had 2 different tests to compare. I want to know if there was much more success, maybe, in one of these, in other words, for example, if it was the same text, the people were more successful. I want to know. I didn't do it yet. I want to know if the existence of certain characteristics, in some of the text, or in some combination of the text, I — I'm not done. And now I'm asked somehow to just, you know, take everything and give it to anybody. I'm not done with my research and that's our research. We have done the work. People who want the research can do research, but its our data. There is also, I must tell you, the issue of — it's very much seems from the things that we are asked to do and from the complaints that there is an attempt to misinterpret our data. This business with how many people contributed to the bad results is that a

(which hardly come up to the usually stated ideals of science),[90] one can be forgiven for suspecting that the distribution data would show

> clear attempt to misinterpret the data and I must admit that I am not, you know, jumping up and down with joy to help to misinterpret my results. I don't think that the people who asked for this data are necessarily, have the necessary expertise. Nobody has ever asked me for data in the past except for colleagues who want to initiate a joint project on the same thing — myself with others. They don't just go around asking for data. If I don't agree, I go ahead and I work. I do my tests. I do my research. And there's also finally, and I understand that this perhaps is the easiest thing to address the issues of confidentiality, if I am ever to release data, I have to go through the dramatic — something I must do, mark out the names. It's handwriting recognition, okay? I'll — We'll have to retype the whole thing. Therefore, I don't have any legal or ethical obligation to do this. The thing for which this is asked of me is to misinterpret it, by you know this contribution of different people and I don't want to do it.

Transcript of Proceedings at 261-63, Estate of John E. Acuff v. O'linger, No. 6064 (Tenn. Ch. Ct. Mar. 25, 1998) (transcript accurately reproduced).

Apparently Saks is regarded by Kam as not having the necessary expertise, since he has failed to respond to numerous requests from Saks to share his data. Interestingly, Dr. Saks is amenable to a joint research project, and has suggested this to Kam in writing, but Kam did not respond to the invitation.

[90] One ideal of science is that data should be shared with the scientific community:

> If all science were conducted according to an ideal, referred to by Robert Merton as the "ethos of science," then scientific findings would be made available to the entire community. Since the purpose of this availability is to allow others to assess the merits of the research, the need for careful description of study procedures is implicit. We believe that, in addition, the availability of the data for scrutiny and reanalysis should be part of the presentation of results. In the past, among the best investigators and with a journal practice open to extensive description, providing data was an honored tradition. Cavendish's classic paper on the density of the earth is a prime example.
>
> Scientific inquiry must be open, and the sharing of data serves to make it so. Disputes among scientists are common; without the availability of data, the diversity of analyses and conclusions is inhibited, and scientific understanding and progress are impeded.

Report of the Committee on National Statistics, *in* SHARING RESEARCH DATA 9-10 (Stephen E. Fienberg, Margaret E. Martin, & Miron L. Straf eds., 1985) (citations omitted). The reasons given by Kam, which are set forth in *supra* note 89, for not sharing data are all dealt with in that volume and found wanting. Kam's claim to a proprietary interest in publicly funded research data is interesting but ultimately unpersuasive. Most government research grants *mandate* data sharing, and even if the FBI does not, public policy demands that publicly funded research data be available to the public unless there is some other good reason to withhold it. "Certainly data collected by government agencies, to the extent that questions of confidentiality and national interest are not present, should be readily and promptly available for research applications. *The same rule should be followed for data collections commissioned for purposes of public policy and for performance evaluation.*" Jerome M. Clubb, Erik W. Austin, Carolyn L. Geda, & Michael W. Traugott, *Sharing Research Data in the Social Sciences*, *in* SHARING RESEARCH DATA, *supra*, at 74 (emphasis added). In the terms of Recommendation 3 of the Committee on National Statistics: "Data relevant to bimodality for the performance of the non-experts and poor performance by a significant cluster of document examiners.

We hope it is now apparent why Kam's conclusion that the results of Kam II "laid to rest . . . the debate over whether professional document examiners . . . possess a skill that is absent in the general population" is more than a little overstated. A more proper conclusion would be that under test conditions not replicating actual practice in significant ways, and as to an undefined range of subtasks not including many of the most important ones of actual forensic practice,[91] the aggregate average performance of document examiners at avoiding false positives was better than that of non-experts, but that this may be an artifact of the varying incentive and disincentive regimes applying to the two groups, perhaps compounded by other factors. More importantly, there is nothing to show that the experts' relative advantage in avoiding false positives when performing on a known test, even if real, is robust enough to persist in the face of expectancy and suggestion effects (properly) excluded from the test, but present in normal practice. This question can only be answered by tests specifically directed toward this issue, or a regime of blind proficiency testing.

C. Professor Moenssens, the Supreme Court, and the Label "Science"

In *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, the Supreme Court recognized the general problem of invalidated expertise in the courtroom and the special problem of expertise representing itself as "science." That special problem springs from the likelihood that juries may be particularly ill-equipped to judge the weight to be given to testimony allegedly derived from "science," and in addition, they may be especially prone to overvalue evidence carrying that label.[92]

---

public policy should be shared as quickly and widely as possible." Report of the Committee on National Statistics, *supra*, at 27.

Clubb questions whether the researchers' usual time-bound monopoly on first analysis should apply to such data, but even if it does, the norm in the scientific community is a duration of one to two years from the initial data collection. *See* Clubb et al., *supra*, at 74. The time is now up for Kam's data. In addition, the fear of "misinterpretation" by "unqualified" investigators is not a valid reason to refuse to share data even though such fear may be inevitable. *See id.* at 57-58; Terry E. Hedrick, *Justifications for and Obstacles to Data Sharing*, in SHARING RESEARCH DATA, *supra*, at 133, 137.

Finally, the very notion of confidentiality in regard to such public interest data on competency and performance of public employees or forensic experts is somewhat problematic, but Kam himself admits it is easily dealt with by redaction.

Where now the "true scientist?"

[91] *See supra* notes 73, 85.

[92] Professor Moenssens has recognized the problem in past writings:

This at least is the law's traditional concern. The Court's opinion makes clear that under Federal Rule of Evidence 702, the trial judge has a general gatekeeping function to evaluate the minimum acceptable dependability of all proffered expert testimony, regardless of its "novelty."[93] The *Daubert* Court attempts to define serious standards for discharging that threshold gatekeeping function in regard to "scientific" evidence, without attempting to map the contours of expertise in general, without giving much guidance on the proper standards for "non-science" expertise, or without stating whether varying threshold standards may be appropriate to different subsets of the category of "non-science" expertise.

Professor Moenssens disagrees with the concept of science he finds in *Daubert*. He believes that *Daubert*'s explicit standard should be limited to what he calls "the 'experimental science' model of Sir Isaac Newton."[94] Most areas of clinical practice, including handwriting identification, Moenssens would dub "science" of another stripe,[95] and he suggests another set of threshold criteria to be applied to them, which he claims will be an appropriate proxy for the validity requirements articulated in *Daubert*.[96]

To be sure, the *Daubert* opinion is not the most sophisticated document ever drafted on the modern concept of science. However, the general notion it reflects is reasonably clear. The *Daubert* Court

---

> [J]uries have been conditioned by the novels they read and the television programs they watch to believe that science can do anything and that scientific evidence is always accurate. When experts testify in a criminal case, the jury frequently perceives that the testimony's value or reliability is far greater than the underlying principles or techniques would justify.

Andre Moenssens, *Admissibility of Scientific Evidence — An Alternative to the* Frye *Rule*, 25 WM. & MARY L. REV 545, 567 (1984) [hereinafter Moenssens, *Admissibility*].

93 "[W]e do not read the requirements of Rule 702 to apply specially or exclusively to unconventional evidence. Of course, well-established propositions are less likely to be challenged than those that are novel, and they are more handily defended." *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 593 n.11 (1993) (unanimous as to that part of decision). Moenssens seems strangely unaware of this language. He writes:

> The case did suggest that the trial judge act as a gatekeeper to prevent the admission of 'unreliable' evidence, but whether such screening function is to be exercised in every case and for all types of expert opinion testimony, or applies only to evidence with which the courts have had little occasion to become familiar (i.e. "novel" scientific evidence) is by no means clear.

Moenssens, *Post*-Daubert *World*, *supra* note 1, at 276. How much clearer does he require it to be?

94 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 252.

95 *See id.* at 310.

96 *See id.* at 291-92.

adopts a notion of science common to every community of conceded scientists, from astronomy to zoology, and denies the label to expertises or claimed expertises derived by methods that do not meet the standards involved in this notion of science. Professor Moenssens' attempt to claim the label "science" for enterprises that do not manifest the minimum characteristics of modern science has no benefit except for the egos of the participants in those enterprises. Further, his attempt not only to win the label "science" for such expertises but also to create alternative, and virtually unfailable, "validity" criteria for them, is a double dip, which would produce very questionable results indeed.

The very term "science" is somewhat problematic. In a general sense, which was its main sense in the 17th century and which is still not entirely unused, the term "science" means any organized and analytical approach to any subject of human knowledge or interest whatsoever. Thus we had "moral science" and still have "political science," "military science," and "library science." In its central modern meaning, however, the term "science" is reserved for what would have been called in earlier times "natural philosophy" and later "natural science." *Daubert* used the term "science" in the latter sense, as Moenssens is well aware.

Science in this sense is not a quality, but an enterprise with a restricted subject matter and a distinctive community methodology. The enterprise of science deals with the generation of a certain kind of human knowledge in a certain way. The kind of knowledge that comprises the realm and object of scientific inquiry is knowledge about the universe accessible to the senses. The way the enterprise of science generates knowledge about the world accessible to the senses, the typical structures of knowledge that this enterprise yields, and the principles by which such knowledge is taken to be validated, are, perhaps somewhat surprisingly, not completely worked out in a way subject to total agreement either among its practitioners or among those who make the examination of the enterprise their own subject of inquiry (that is, philosophers and historians of science). Nevertheless, the general characteristics of the enterprise are fairly well understood.

Any sophisticated account[97] of the scientific enterprise must give due regard to the role of considering a wide range of reported sense

---

[97] We do not claim that the short account in the text is very sophisticated. Indeed, we have tried to avoid controversy by writing only in the most general terms with few footnotes. Still, the main contours set out here will suffice for our purposes.

observations in laying out the contours of the possible, the role of imagination and esthetic intuition in forming hypotheses about possible worlds consistent with the data regarded, and the role of skeptical formal empirical testing in attempting to falsify the imagined world of the hypothesis. Sense observation of the physical world, perhaps aided by sense-enhancing instruments developed for the purpose, is the foundation of science, but of course disorganized reports of more or less random sense observations do not form part of the process of science. Science requires standardization of the conditions of such reports, and a formal analytical system for their organization. A central condition that must be present is theoretical reproducibility of observation (two observers in the same position could perceive the same thing), with strong favor given to reports of observations that can be practically repeated by multiple observers, with stronger favor given to reports that reflect dimensions of an event that can be measured, and even stronger favor given to dimensions that can be measured quantitatively with some precision.[98] These reports must then be organizable, and in fact be organized, according to some analytic system of similarities and differences that generates hierarchies of categories. In short, the enterprise of science requires an explicit formal and generally well worked out taxonomy.

These taxonomic conditions for the collection and organization of observations are necessary conditions for the enterprise of science, but not sufficient in and of themselves to comprise a science. This is not to say that the descriptive organization of observations is not valuable in itself. Whole branches of human knowledge, including much of history and philosophy, are primarily of this sort. The taxonomic enterprise is not unscientific, but it does not by itself make a science. Science must go beyond descriptive taxonomy of even the most quantified sort to the enterprise of hypothesis and theory creation.

A scientific hypothesis is a statement about inter-relationships between items or categories that is formulated in such a way that it can be subjected to empirical testing. A theory is a set of intercon-

---

[98] We mean to take no position on controversies concerning the point at which apparent precision becomes misleading illusion or the proper role for ways of dealing with various forms of indeterminacy (such as Bayesian probability theory, or the mathematics of fuzzy sets) in any aspect of the scientific enterprise, including taxonomy, hypothesis formation, and theory building. Even in the face of various indices of indeterminacy, whether conceived of as inevitable or not, testability and the requirement of testing remain constants in all approaches.

nected hypotheses of varying generality which account for a wide range of phenomena.[99]

A hypothesis that cannot be subjected to empirical testing is a metaphysical proposition that is by definition not part of science. A hypothesis that is potentially subject to testing but has never been tested is unproven, and cannot be treated as a source of dependable scientific knowledge. A hypothesis that has been subjected to substantial empirical testing and has not been falsified may properly be treated as validated, although all empirical testing is essentially a probabilistic enterprise. Thus, no hypothesis is ever subject to absolute verification, but the nature of both the data and the fit of the hypothesis with other more or less validated[100] principles may push the probability of the truth of the hypothesis so high that it would be crazy to spend much time worrying about the residual probability of falsity for any practical purpose. Such relations may come to be referred to as scientific "laws."

The third major aspect of the scientific enterprise is real-world empirical testing of the truth of the relationships asserted by hypotheses, and the theories that both derive from and generate hypotheses. Like the taxonomic aspect, the empirical testing aspect involves careful observations of phenomena through the senses. However, empirical testing is focused on the observations capable of testing (falsifying) a hypothesis, whereas taxonomically important observations may be made in a much less directed way. The empirical testing enterprise contains a whole constellation of special problems revolving around how to determine what observations and conditions of observation are in fact consistent or inconsistent with a hypothesis, with one of the chief concerns being to structure conditions in such a way so as to avoid what are globally called "artifacts," especially "expectancy effects," one of the most serious dangers dogging the

---

[99] "Facts are the world's data. Theories are structures of ideas that explain and interpret facts." Stephen Jay Gould (1983), *quoted in* ISAAC ASIMOV & JASON A. SCHULMAN, ISAAC ASIMOV'S BOOK OF SCIENCE AND NATURE QUOTATIONS 100 (1988). Some courts mistakenly assume that science deals only with second and third order abstractions, properly labeled theories, which are directed at explaining the cosmos, or at least "the world," and not with more concrete and everyday phenomena. This was the fundamental error of the court in *United States v. Jones*, which appears to have concluded that handwriting experts are not and could never be doing science: "We are quite convinced that handwriting examiners do not concentrate on 'proposing and refining theoretical explanations about the world.'" United States v. Jones, 107 F.3d 1147, 1157 (6th Cir. 1997) (quoting *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 590 (1993)).

[100] *See infra* note 103 (on the tension between validation and falsification).

quest for scientific knowledge.[101] In many areas of science, empirical testing must not only meet such criteria, but new instruments must be developed that will make such observations possible.

The taxonomic, the hypothesis making/theory generation, and the empirical testing aspects are not totally separate enterprises, of course, and every practicing scientist will have familiarity with each process, at least within the confines of his or her area of interest. Each process feeds on the other, with taxonomic systems providing material that theories must take into account, theories providing predictions (hypotheses) to which testing must be directed, and testing observations providing new data, which must be absorbed into taxonomies. It is interesting to note that individual practitioners within the enterprise of science may show marked predilections for one function or another. People whose strength is hypothesis generation/theory building may, for reasons of esthetics or ego, display quite "unscientific" commitment to the validity of their hypotheses in advance of testing (which may sometimes bear them out and sometimes shatter them). In fact, it may even be that such an "unscientific" emotional commitment is necessary to drive some such people through the act of creative imagination necessary for the hypothesis generation. At the other extreme, people whose careers are devoted primarily to testing other people's hypotheses can come to be regarded, rightly or wrongly, as unimaginative skeptics driven by a vision of themselves as the reality police of the enterprise of science. Finally, taxonomists tend to be (unfairly) regarded by all sides as, to borrow the phrase of Samuel Johnson describing lexicographers, "harmless drudges."[102] Uncommonly, a single individual may combine substantial talents in all areas. For example, Darwin was a first-class primary observer and taxonomist as well as theoretician, who was quite sensitive to the need for testing. Others may not. Einstein made few primary observations and was not particularly interested in the dirty details of confirmatory (or disconfirmatory)[103] testing

---

[101] *See generally* ROBERT ROSENTHAL, EXPERIMENTER EFFECTS IN BEHAVIORAL RESEARCH (1976). Accordingly, real science aims to structure its procedures in ways that will maximize the contribution of evidence and minimize the contribution of bias to data gathering and inference drawing. For example, consider the widespread use of double-blind research designs.

[102] SAMUEL JOHNSON, DICTIONARY OF THE ENGLISH LANGUAGE (6th ed. 1785).

[103] Here we must address the main technical conflict between Karl Popper and his critics, which results from Popper's attempt to solve the Problem of Induction. *See* COLIN HOWSON & PETER URBACH, SCIENTIFIC REASONING: THE BAYESIAN APPROACH ch.1 (1989). Popper's claim is that one can disprove a hypothesis in a way significantly different than one can prove it. Basically, as a matter of logic, ignoring second order problems of knowledge, a series of confirming instances, no matter how

(though he of course recognized the necessity for it). Nevertheless, for an enterprise and its products to qualify as a science, its group practice must manifest all three functions in a balanced, ongoing, and dynamic way, even if its individual practitioners do not always do so.

It is probably not very controversial to say that, within the realm of factual inquiry, the properly tested products of the enterprise of science, where relevant, are the most dependable sources of factual information available to human beings. This does not mean, however, first, that valid science-generated information will always be available on the fact issues the law cares about or, second, that other non-science sources of fact information are not dependable enough for the practical purposes of the law in many contexts. However, the special preeminence of science in the realm of fact more or less dictates that anything being labeled science in the courtroom ought to meet the validity criteria of proper science. Exegesis of particular language aside, this we take to be a central precept of *Daubert*.

Professor Moenssens objects to the general application of these criteria, and raises what he implies is a parade of horribles that would flow from such a standard, including the exclusion of many "soft science" subject matters such as psychological syndrome evidence, etc. (of which, he claims, handwriting identification is one).[104]

There are numerous problems with Moenssens' position. First, most of the things he includes as "soft science," such as the conclusions of forensic pathologists and various examples of psychological "syndrome" evidence, are clearly "part of science" in the general sense of being the products of practitioners who claim to be applying scientific knowledge developed elsewhere in the general enterprise,

---

long, does not establish a proposition, because the next instance encountered may be the contradictory instance that falsifies the proposition, whereas one falsifying instance falsifies. This is the basis for Popper's emphasis on falsifiability. However, in the real world, one cannot ignore second order problems of knowledge. Our knowledge of the reality of disconfirming instances is as probabilistic and subject to error as our knowledge of confirming instances, which undermines the logical difference between confirmation and falsification and reduces both to probabilistic enterprises. Whether or not this insight totally destroys the special place of falsification is a topic of debate, a debate which was unlikely to have been well understood by the Supreme Court when it invoked Popper and which does not make any practical difference in real world applications. Even committed Popperians will operationally treat propositions as "confirmed" when the series of uncontradicted confirmatory instances gets so high that worrying about the remote probability of falsification seems a waste of time. As Moenssens does not seem to understand, none of the debates surrounding Popper's approach have much, if anything, to do with the issues concerning expert testimony addressed in *Daubert*. *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 286-87

[104] *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 294.

or who claim to have developed such knowledge themselves. These experts are, in short, practitioners who are part of an ongoing scientific enterprise. That many of them, or perhaps most in some areas listed by Moenssens, have confused hypothesis with fact,[105] merely illustrates the usual operation by which the group makes progress even though many credentialed members of the group may in the interim act "unscientifically." This phenomenon certainly does not justify creating a new category for such experts, much less applying any relaxed standards of validity to them. If such practitioners are not allowed to testify because they fail the minimum validity criteria proper to a scientific enterprise, that is not necessarily a bad thing. And if we allow them to testify in those circumstances, perhaps we had better work out a better account of what we are about when we do so. Certainly, allowing practitioners unable to meet minimum validity criteria to give testimony assertedly relevant to actus reus or to identity issues[106] in criminal cases is among the least justifiable thing we can do if we mean what we say about the presumption of innocence and the principle of rational fact finding.[107]

---

[105] It is curious to us that Professor Moenssens can celebrate his own justified skepticism toward such things as earmark identification, *see* Moenssens, *Post*-Daubert *World, supra* note 1, at 293-94, and aspects of syndrome, footmark, and bitemark identification, *id.* at 292 & n.109, in an article which displays such an unquestioning commitment to all applications of Osbornian handwriting identification in spite of the weakness of the evidence concerning it. In this regard, it is instructive to compare his attitude toward handwriting identification with the positions he has previously expressed toward another asserted identification-by-comparison technique, voiceprints. In 1984, after reviewing the literature on voiceprints (which presumably he considered himself qualified to do), he concluded that voice spectrography "produces results based on an unproven assumption." Moenssens, *Admissibility, supra* note 92, at 557. Moenssens further noted that:

> The theory of voice uniqueness, however, which the scientific community should have subjected to a searching stage-two inquiry, never has been proven by empirical evidence. As long as the theory remains a postulate, and not a proven fact, the technique of comparing voice spectrograms cannot establish the identity of a speaker with any relative degree of certainty. Upon what basis then could a court find that voiceprint identification had been generally accepted?

Id.

[106] On the special place of such issues in the establishment of truly factual and non-normative innocence, see generally D. Michael Risinger, *John Henry Wigmore, Johnny Lynn Old Chief and "Legitimate Moral Force:" Keeping the Courtroom Safe for Heartstrings and Gore*, 49 HASTINGS L. J. (forthcoming 1999).

[107] Moenssens himself would agree with the need for reliability of expert methodology:

> A conservative test — one that will not permit the prosecution or the defense to use evidence based on tests of unproven reliability — is especially commendable in criminal cases, in which the technical inquiry is directed toward resolving the ultimate issue of guilt. When scientific

However, handwriting identification (the main focus of *Post-Daubert World*) doesn't even make it into the category of "science" in any tenable modern sense.[108] No members of the handwriting identification community are rewarded for doing empirical testing and for examining the claims of the enterprise skeptically. To apply the label "science" to the enterprise in general or to the practice of individual practitioners in particular would deprive the term of any defensible meaning.[109] Note that this does not mean that, *a priori*, it should be

---

> tests purport to identify the defendant as the guilty party, courts must screen out those techniques that do not invariably point to the guilty party. Otherwise, the jury might be deceived.

Moenssens, *Admissibility*, *supra* note 92, at 546 (footnotes omitted). Moenssens further noted that:

> In criminal cases, where an individual's freedom is at stake, courts certainly ought to be very cautious in admitting evidence based upon insufficiently tested or verified premises, especially when the evidence seeks to establish the ultimate issue in the case — the identification of the accused as the perpetrator of the offense.

ANDRE A. MOENSSENS, JAMES E. STARRS, CAROL E. HENDERSON, & FRED E. INBAU, SCIENTIFIC EVIDENCE IN CIVIL AND CRIMINAL CASES 18 (4th ed. 1995).

[108] Compare the basics of the scientific method to the comments of leading questioned document examiners on a panel discussing the controversy over the scientific shortcomings of QDEs. Robert Muehlberger: "We are not research scientists. We are practitioners and have a caseload." Duayne Dillon, referring to the inability of QDE practitioners to convince others that their techniques are valid and reliable: "Only we who do it can know that what we say about it is true." Mary Wenderoth Kelly, referring to the skeptical comments of Risinger or Denbeaux on the same panel discussion: "He went after our weak point: no data." Contemporaneous notes of Michael J. Saks from Annual Meeting of AAFS, Nashville, Tenn, (Feb. 19-24, 1996) (on file with the *Seton Hall Law Review*).

[109] There are three different ways that QDEs could attempt to find scientific (as opposed to merely rhetorical) support for their claims. These are the following:

1. The applied science model — theoretical extrapolation version. If it could be shown that QDEs are applying principles that have been well established (empirically tested) in basic research, that provides a basis for believing there would be validity in the extrapolation of that basic science knowledge to QDEs' clinical practice. For example, compare forensic chemists, who borrow the basic knowledge of chemistry and apply it to forensic chemistry problems. (Of course, the correctness of the application and the accuracy of the results produced still need to be examined.) But, of course, for handwriting identification there is no body of basic research knowledge from which QDEs can borrow. "*If there is no science, there can be no forensic science.*" John Thornton, *The General Assumptions and Rationale of Forensic Identification: Substitution of Intuition or Experience for Defensible Scientific Fact*, *in* MODERN SCIENTIFIC EVIDENCE, *supra* note 40, § 20-5.5.

2. The applied science model — statistical modeling version. The fundamental assumptions of QDEs are similar to those of DNA typing: intense variability among a large number of attributes. Accordingly, QDEs could undertake to do what has been done in DNA typing: Obtain counts of frequencies of various writing attributes in the population, and use those actually to calculate the probability of a coincidental erroneous match in the case at bar (rather than making wholly intuitive and subjective judgments). *See generally* Michael J. Saks & Jonathan J. Koehler, *What DNA*

impossible for any handwriting identification expert ever to testify. Instead, this means that handwriting identification must be treated as a claimed non-scientific clinical skill. For such asserted but untested expertise, from dowsing to wine tasting, what is lacking in internal scientific validation must be supplied by external scientific validation, treating the practitioners not as scientists but as instruments, and testing their accuracy under a variety of conditions.[110] Only by requiring such external scientific validation of clinical expertise can the law obtain what it claims it is after — information that can be rationally helpful because there is reason to believe it dependable beyond the self-confidence of its practitioners.

Which brings us to our final point. Even if it were defensible to create a subset of "clinical science" for purposes of determining admissibility, and even if handwriting identification were in such a category, the criteria that Moenssens has proposed for determining admissibility are so tenuously related to validity and Daubertian "reliability" as to be virtually no test at all. Here they are, exactly as Moenssens set them out:

"Fingerprinting" Can Teach the Law About the Rest of Forensic Science, 13 CARDOZO L. REV. 361 (1991). While small beginnings to such a science have been attempted, it has had no application by QDEs. *See* sources collected in *Exorcism*, *supra* note 2, at 739 n.31.

3. The black-box model. Taking QDEs and their practice as it has existed for the past century or so, one still can easily bring the scientific method to bear simply by treating the QDEs and the process of doing whatever it is they claim to be doing as a black box, and measuring the inputs and outputs. That is, QDEs can systematically be presented with writings of known origin, and their decisions about authorship can be compared to what is known to be true of those writings. By testing different kinds of QDEs with various different kinds of stimulus writings, under different types of testing conditions, using different methods of examination, one could eventually map the extent of special abilities and limitations of different types of QDEs, and of different individual QDEs, in examining different types of writing, using different comparison methods, under different types of conditions. Moreover, such research would provide empirical answers to many of the assertions about QDEs, and differences among QDEs, that are sprinkled throughout Moenssens' article, with nothing but his (and the QDEs) *ipse dixit* for support.

[110] We have made this point elsewhere. *See* Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 40-41.

SUGGESTED CRITERIA FOR TESTING "RELIABILITY"

> (1) A professional literature exists in a field that describes the accepted processes by which the aims of the field can be reliably realized.[111]
>
> (2) There exist professional societies and associations, open to the membership of individuals whose qualifications have been examined, that engage in continuing education, publication and research pursuits.
>
> (3) The discipline possesses recommended courses of study and training for becoming competent to work as a professional in the field, and access to the profession is open to all qualified individuals.

---

[111] This apparently refers to the "scientific method of comparison," which Moenssens claims justifies forensic identification independent of any attempt at being quantified statistically in terms of error rates. *See* Moenssens, *Post*-Daubert *World, supra* note 1, at 320. By "scientific method of comparison," therefore, Moenssens apparently means the method recommended by the field, what has been sometimes referred to in the case of Osbornian handwriting identification as the "standard methods," "proper methods," or simply "true methods" by its leading practitioners. *See* authorities cited in *Science and Nonscience, supra* note 2, at 35 n.90. One should understand, therefore, that there is a world of difference between "the scientific method" as that phrase is usually used in science, and Moenssens' "scientific methods of comparison." In regard to handwriting, Moenssens actually describes the "scientific method of comparison" as follows:

> In *United States v. Velasquez*, 64 F.3d. 844, 846 n.3 (3rd Cir. 1995), the government expert described the procedures that she and other experts in handwriting analysis employ: (1) the expert determines whether a questioned document contains a sufficient amount of writing and enough individual characteristics to permit identification; (2) if the document is identifiable, the expert examines the submitted handwriting specimens in the same manner; (3) if both contain a sufficient number of identifiable features the two documents are compared to determine whether the individual features are different or the same; (4) the expert weighs the evidence, considering both similarities and differences in the writings examined and determines whether or not they are by the same writer.

Moenssens, *Post*-Daubert *World, supra* note 1, at 320 n.306. Here is what the *Starzecpyzel* court thought of this "scientific method of comparison," as explained by a prominent QDE expert witness, Mary Wenderoth Kelly: "[H]ow FDE's [Forensic Document Examiners] might accomplish this [identification] was unclear to the Court before the hearing, and largely remains so after the hearing." United States v. Starzecpyzel, 880 F. Supp. 1027, 1032 (S.D.N.Y. 1995). "Both the Defendants and the Court sought elucidation on this issue, but met with little success." *Id.* Based on Moenssens' summary, it is not hard to understand why: Document identification is an entirely subjective process. Kam, writing on the same point, said: "It is very likely that many examiner decisions and associations are difficult to verbalize, and that some verbal explanations are post factum recreations of the reasoning process." Kam I, *supra* note 8, at 12.

> (4) The education and training to achieve basic proficiency in the discipline is supervised by professionals in the discipline who have gained some degree of prominence in their field, and who can impart knowledge to students of the discipline as they are supervising the development of the students' basic skills.
>
> (5) There exists a program of supervised practical or clinical experience under the guidance of experts who have gained some degree of recognition in the discipline.
>
> (6) A person offering to testify as an "expert" in the discipline has demonstrated the aptitude and proficiency expected of working professionals in the field by an examination board, a sponsoring professional association, or by completing other credentialing procedures.
>
> (7) There exists the ability to retest the same evidential materials by an expert for the opposing side, or for the participation of such an expert in testing that is destructive in nature.[112]

Moenssens then states that this list is "neither essential nor exclusive," but that it will "screen out the type of 'forensic' expert opinion that is truly 'unreliable.'"[113]

Even were these criteria absolutely required, however, there is nothing in them that will screen out the "unreliable." Meeting these criteria merely establishes the existence of a self-believing guild with a modicum of organizational skill. It might surprise Professor Moenssens to realize that by his "test" many of the "graphologists" would be able to testify, since their guild structure mirrors that of his favorite group, the American Society of Questioned Document Examiners (ASQDE), or can be easily modified to do so.[114] Indeed, his "test" seems to have been generated by asking "what criteria can the ASQDE members satisfy," and backing the list out from there. Note that there is no requirement in Moenssens' criteria of true profi-

---

[112] Moenssens, *Post-*Daubert *World, supra* note 1, at 291-92.

[113] *Id.* at 292.

[114] The facts underlying this assertion are provided by Moenssens himself. Indeed, one of his complaints is that graphologist groups have tried to mirror the ASQDE. *See* Moenssens, *Post-*Daubert *World, supra* note 1, at 269-70. There seems to be little reason why the Independent Association of Questioned Document Examiners or the Association of Questioned Document Examiners, described by Moenssens in his appendix, could not satisfy all of Moenssens' listed criteria with minor adjustments to their practices and by-laws. At least one of the two organizations publishes a journal, the *Journal of Questioned Document Examination,* which is older than the Osbornian *International Journal of Forensic Document Examiners,* which only began publication in 1995.

ciency testing, only a requirement that proficiency in doctrine and accepted procedure be demonstrated to a board of "true believers," independent of validity.[115] The weaknesses of The American Board of Forensic Document Examiners certification testing procedure are well documented[116] and show the weakness of Moenssens' criteria as drafted. Moreover, one might think that the requirement that some members of the enterprise be engaged in "educational and publishing pursuits" under criterion 2 might provide some scientific context, until one recalls what weak anecdotal material Moenssens is willing to count as "research." In short, Moenssens' criteria are nothing more than a version of the *Frye* test unhooked from any standards of science, combined with the most narrow and formalistic notion of what constitutes a proper community of acceptance.[117] They go counter to the direction of both the Supreme Court and the Advisory Committee on the Rules of Evidence in seeking true reliability criteria for application in contexts beyond the realm of "science."[118]

---

[115] Elsewhere, Moenssens has written that "[m]ost of the witnesses who testify as experts for the prosecution are not truly scientists, but better fit the label 'technicians' . . . ." Andre A. Moenssens, *Novel Scientific Evidence in Civil and Criminal Cases: Some Words of Caution*, 84 J. CRIM. L. & CRIMINOLOGY 1, 5 (1993) [hereinafter Moenssens, *Novel Scientific Evidence*]. Such technicians "are not however the kind of experts that lawyers need in an attempt to make new law by being among the first to introduce novel expert testimony in court." *Id.* Need we add that this applied equally to old but unvalidated principles and techniques. If one takes Moenssens' criteria at face value, restrictive organizations of astrologers could readily be created that would qualify their members to testify.

[116] As we pointed out in *Science and Nonscience*:

> [T]he certification testing program of . . . the American Board of Forensic Document Examiners . . . described by Mary Wenderoth Kelly (a member of that certifying board) in her *Starzecpyzel* testimony, leaves a lot to be desired. While there is a short multiple choice test to measure knowledge of handwriting identification doctrine, the heart of the examination is based on the administration of five of only seven or eight test problems, only two or three of which involve handwriting identification. The same problems are used year after year on an honor system where they are sent to the candidates for certification through their teaching mentor, and left with them for a month unsupervised before the answers are returned.

Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 38 n.94.

[117] *See* Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence:* Frye v. United States *a Half-Century Later*, 80 COLUM. L. REV. 1197, 1209-11 (1980) (discussing, among various shortcomings of the *Frye* test, the fact that the more narrowly one defines the expert community the more general acceptance will appear to exist, and the more broadly one defines the community the less consensus will appear).

[118] The proposed amendment to FED. R. EVID. 702, drafted by the Advisory Committee on the Rules of Evidence to the Committee on Rules of Practice and Procedure of the Judicial Conference of the United States, reads:

> If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue,

Finally, Moenssens' criteria are focused on the seeming legitimacy of groups of people, not on the validity of components of their knowledge or techniques. Thus, surgeons would surely pass Moenssens' test, and all of their favorite surgical techniques would be "recognized" and "respected" (to use his terms). But, as we previously noted,[119] a review of systematic empirical research on many of the most commonly employed surgical procedures discovered that only about one-third of them worked as well as believed, another third were ineffective, and the final third actually did more harm than good. Surgeons realize that systematic empirical testing is their best guide to what works and what does not, not the sorts of factors upon which Moenssens would have the courts rely. Thanks to their empirical research, surgeons can continually improve the quality of their work and over time shed their errors.

Perhaps the worst effect of Moenssens' proposal, but the effect that he apparently most hopes for, is that pesky skeptics who are not guild members would be barred from testifying about potential weaknesses in the expertise. On this point, as on other points in his article, Moenssens is not entirely consistent, since he is willing to celebrate non-guild members like Kam in whose work he can find support.[120] But it is clear that when it comes to evaluating the dependability of handwriting identification, Moenssens finds consistency to be an inconvenient hobgoblin of small minds.[121] Perhaps the

> a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, provided that (1) the testimony is sufficiently based upon reliable facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.

COMMITTEE ON RULES OF PRACTICE AND PROCEDURE, JUDICIAL CONFERENCE OF THE U.S. PRELIMINARY DRAFT OF PROPOSED AMENDMENTS TO THE FEDERAL RULES OF CIVIL PROCEDURE AND EVIDENCE (August 1998) (circulated for comments). This amendment appears designed to move the courts further along the path of validity and clarity that the Supreme Court blazed in *Daubert*, and further away from the looseness recommended by Moenssens.

119 *See supra* note 34 and accompanying text.

120 Moenssens has some very slippery definitions of what QDE science is and who can do it. He says that QDEs are not scientists but technicians and as such cannot be expected to do empirical research, to know the literature, or to be able to talk about it in court — tasks to be done by those who know about empirical research. But when we undertook that evaluation (by looking at the empirical research literature, or lack thereof), he says that people like us are not qualified to do it, because we are not trained QDEs (though he makes a special exception for Kam). What is one to make of this? Moenssens seems to be saying: You can't ask QDEs to justify themselves. And non-QDEs cannot inquire into the empirical basis of QDEs' claims, except when they produce studies that appear to be supportive.

121 Consider the positions taken in *Post*-Daubert *World* with Moenssens' many pre-

current fever will subside and he will return to the more defensible positions of his past. We can only wish him a speedy recovery.

PART II

Introduction

At various points in Professor Moenssens article we are told the following about ourselves and our work: We are "self styled evidentiary law reformers,"[122] "legal activists"[123] who have descended "from the perch of academic inquiry into partisan advocacy";[124] our work "lack[s] intrinsic worth,"[125] "suffer[s] from serious research deficiencies,"[126] is "highly questionable research,"[127] is marked by "superficiality,"[128] and is "highly deficient in scholarship and based upon an ignorance of the underlying factual basis."[129] Our "poor scholarship"[130] is properly characterized as "junk scholarship,"[131] "inadequately researched from a technical as well as a legal standpoint";[132] we failed in our research,[133] and "the author's serious deficiency in legal research"[134] reflects "innumerable errors"[135] using "inaccurate, inappropriate, out-of-date citations"[136] and "citing authorities for propositions which they do not support."[137] Moenssens claims that our writing has "distortions, omissions and many other shortcomings"[138] and "many inaccuracies and slanted statements,"[139] is "poorly documented"[140] and in general is "a lot of aca-

---

vious positions noted throughout this article, particularly *supra* notes 35, 92, 105, and 107, and *infra* notes 147 and 283. Sometimes internal inconsistencies appear in *Post*-Daubert *World* itself. *See supra* notes 35, 105, and 114 and *infra* notes 164, 236, and 239.

[122] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 266.

[123] *Id.*

[124] *Id.* at 274.

[125] *Id.* at 256.

[126] *Id.*

[127] *Id.* at 307.

[128] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 254 n.7.

[129] *Id.* at 299.

[130] *Id.*

[131] *Id.* at 330 n.348.

[132] *Id.* at 311.

[133] *Id.* at 302 n.224.

[134] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 304.

[135] *Id.*

[136] *Id.*

[137] *Id.*

[138] *Id.* at 309-10.

[139] *Id.* at 306.

demic hogwash,"[141] "a pseudo-study"[142] "ignoring, denying or suppressing inconvenient evidence";[143] we are "appallingly ill-informed and misinformed"[144] and our conclusions are "unwarranted, untenable and erroneous."[145]

Needless to say, we do not agree with these characterizations, and we will deal with the asserted bases for them (to the extent there are any) at some length in this Part. However, we think it necessary to put the rhetorical excesses of Professor Moenssens' article on the table at the outset, if for no other reason than to give the reader unfamiliar with the piece some exposure to its peculiar viciousness. In Part I we characterized Professor Moenssens' failure to follow the more judicious principles of his previous writings as a fever, but in fact it appears to be more in the nature of a fit. One has to wonder at the forces that have precipitated all this vitriol. Indeed, we have a few hunches.

It seems that we have inadvertently gored a whole herd of Professor Moenssens' oxen. Professor Moenssens entered the forensic science field in the early 1950s in his native Belgium under the tutelage of "the late Major (ret.) Georges H. Defawe."[146] He thus apparently started as one of those "technicians" (his word)[147] who were basically individuals trained in some forensic technique without substantial formal science background. Such technicians constituted the vast bulk of forensic science professionals until quite recently. In addition, Major Defawe was an Osbornian document examiner, and Moenssens himself was trained as a questioned document examiner (QDE) for two years before deciding to specialize in fingerprint identification.[148] He has maintained strong ties to the orthodox QDE community through the years, and was incorporator of the ASQDE when it became a non-profit corporation in 1972.[149] In 1997 he was given a special award by the Document Section of the American Academy of Forensic Sciences.[150]

---

140 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 310.

141 *Id.* at 309 n.260.

142 *Id.* at 311.

143 *Id.* at 309.

144 *Id.* at 330.

145 *Id.* at 330 n.347.

146 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 308 n.253.

147 "Most of the witnesses who testify as experts for the prosecution are not truly scientists, but better fit the label 'technicians.'" Moenssens, *Novel Scientific Evidence*, *supra* note 115, at 5.

148 *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 253.

149 *See id.* at 258 n.22.

150 The award was presented to him by Robert J. Meuhlberg, one of the QDEs

In addition to his emotional ties to the QDE community, Moenssens obtained an LL.M. from Northwestern Law School in 1967, where he came under the personal mentorship of Fred Inbau, who in turn was a protégé of John Henry Wigmore. The almost filial relationship of Inbau to Wigmore and Moenssens to Inbau[151] appears to have provided another strong impetus driving Moenssens' rhetoric while clouding his judgment. Indeed, it almost seems that our worst sin, in Moenssens' view, is a kind of *lese majeste*, an insufficient manifestation of reverence for Wigmore, Inbau, Albert Osborn, and the average (Osbornian) toiler in the field of document examination.

The single most extreme example of our irreverence, in Moenssens' eyes, seems to be our passing characterization of Wigmore as the "800 pound gorilla of American Evidence law."[152] Since Moenssens repeats this phrase with undisguised horror three different times in his article,[153] let us make one thing clear immediately: We do not view this characterization as "derogatory" or even disrespectful (though it is concededly lacking in reverence). Each of us would be gleeful to know that someone would refer to us as the 800 pound gorilla of anything fifty years after our death. The point of invoking this old joke was to emphasize that with position and authority, however well earned, comes the power to throw one's weight around unreflectively and perhaps the temptation to do so, which is not always resisted. We believe Wigmore did this from time to time, and that his uncritical support of Albert Osborn's principles of handwriting identification was one of those times. Wigmore was a giant figure of American jurisprudence, no doubt about it, but he was in many ways a complicated and at times difficult character. We do not believe that anyone can read what we actually wrote about Wigmore without concluding that we attempted to recognize both his strengths (such as his extreme general intelligence, his breadth of interests, and his gi-

Moenssens credits in his prefatory footnote and then an officer of the Questioned Documents sections of both the AAFS and the ASQDE. The award speaks of Moenssens' "unwavering support of the questioned document community." Contemporaneous notes of Michael J. Saks from annual meeting of AAFS, New York, N.Y. (Feb. 17-22 , 1997) (on file with author).

[151] Inbau went to Northwestern Law School for an LL.M. after graduating from Tulane Law School. After receiving his LL.M. in 1933, he was hired by Wigmore to work at the Northwestern University Scientific Crime Detection Laboratory, and later joined the Northwestern faculty in 1936. *See* AALS DIRECTORY OF LAW TEACHERS, (1995-96). Moenssens became Inbau's co-author very soon after his graduation from Northwestern in 1967.

[152] Risinger et al., *Exorcism*, *supra* note 2, at 768.

[153] *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 300 n.222, 302 nn.226 & 229.

gantic capacity for work),[154] and his weaknesses (such as his uncritical devotion to "scientism").[155] This approach apparently strikes Moenssens as inappropriately irreverent.

Similar observations could be made in regard to our treatment of Albert Osborn, who, according to Moenssens, was the author of a "pioneering treatise, a work that still stands revered,"[156] and a "disparaged and maligned [by us, presumably] giant of scientific document examination."[157] Osborn was, as we said, clearly an ambitious man of exceptional intelligence. In *Exorcism*, we tried to give a feel for how such a man of his time and place could become personally committed to the unverified theories he put forth and promoted. Perhaps the harshest words we used to characterize him were "vanity" and "arrogance,"[158] but we do not think one can read his voluminous works, or transcripts of his testimony in the Hauptmann trial, without being reminded of the old saw, "I may have been in error, but I've never been in doubt." Again, this seems to lack the level of reverence Moenssens thinks is appropriate.[159]

We have already dealt with Moenssens' misplaced defensiveness as to the Inbau study and our treatment of it in Part I, and need not repeat ourselves here, except to make two points. First, Moenssens falsely suggests, when giving examples of our "sarcasm," that we disrespected Inbau by referring to "'prisoners,' 'second graders,' and 'calligraphers'"[160] when discussing his study. Second, Moenssens suggests that we were inappropriately harsh when we wrote of Inbau's study, "[W]hat we might infer of the conditions of test administration from some of Inbau's description is hardly encouraging."[161]

---

[154] *See* Risinger et al., *Exorcism, supra* note 2, at 767-68.

[155] *See id.* at 768-69.

[156] Moenssens, *Post*-Daubert *World, supra* note 1, at 257 n.17.

[157] *Id.* at 302 n.224.

[158] *See* Risinger et al., *Exorcism, supra* note 2, at 766.

[159] Moenssens also criticizes us as "demeaning and depreciating" on page 302 and in note 228 for saying that in 1876, "Osborn's mastery of penmanship offered a path off the farm." Moenssens, *Post*-Daubert *World, supra* note 1, at 302 n.228. But this is indubitably true, and hardly disrespectful. As was explained at length in *Exorcism*, drawing from Clark Sellers's laudatory biographical sketch, Osborn had dropped out of college and returned to his father's farm where he lived until he got a job teaching penmanship at a business college in Rochester. *See* Risinger et al., *Exorcism, supra* note 2, at 764-65 (citing Clark Sellers, *Albert Sherman Osborn: Questioned Document Pioneer*, 45 A.B.A. J. 1285 (1959)). Leaving the farm thus is an honored tradition. One of our own fathers sought and found a path off the farm in an analogous way, as have the forebears of millions of Americans.

[160] Moenssens, *Post*-Daubert *World, supra* note 1, at 302 n.226.

[161] *Id.*

As to the first point, the words given had nothing to do with the Inbau study specifically. Here is what we wrote regarding test design in general (the passage from which the "sarcastic" words were drawn):

> Identifying handwriting from stylistic similarities perceived by juxtaposing authentic exemplars with questioned documents presents many of the same problems as eyewitness identification. Some circumstances are so clear that mistakes are unlikely on anyone's part. If one of five *prisoners* must have written a threatening note slid out of a jail cell, four of them have to that point printed like *second graders* every halting document they have ever been know to produce, the fifth is a professional *calligrapher*, and the note is in flowing calligraphy, then identifying the writer would be easy for anyone. On the other hand, some identification tasks may be beyond anyone's skill.[162]

We may at times have been sarcastic, but the above example was clearly not one of those times.

As for the second challenged passage, we do not know how we could have more gently introduced the point, fully documented with quotes from Inbau's own report, that Inbau had shared his test materials and the correct answers with colleagues, who could have, and in at least one instance did (we will assume inadvertently, as Inbau says) destroy the test by sharing this information in advance of the tests with people who turned out to be test subjects. So much for Moenssens' sensitive protectiveness toward Inbau. Filial defensiveness is understandable, and perhaps even admirable in some contexts, but it does not make for balanced scholarly evaluation.

As to our attitudes toward the average document examiner, we have a great amount of sympathy for them. They have invested heavily in the reality of what they do, and questions about its limits are understandably threatening and painful. However, we also have sympathy for those whose fates turn on their testimony to one degree or another. Moenssens characterizes our "tone" as having been "sarcastic," "scornful," and "demeaning and depreciating."[163] We believe he is displaying undue sensitivity. We wrote to drive our points home effectively in the face of an established order and received wisdom. If we made the keepers of that order and received wisdom uncomfortable, that was an unfortunate side effect of a necessary process.

---

[162] Risinger et. al., *Exorcism, supra* note 2, at 742 (emphasis added).

[163] Moenssens, *Post-*Daubert *World, supra* note 1, at 302.

Whatever may explain the intensity of the Moenssens rhetoric, the form of his article is yet another puzzle, and one that complicates the framing of this response. First, it is long (80 pages). Second, it is oddly fashioned, with many internal contradictions, repetitions, and infelicities.[164] The article gives the appearance to us of a pastiche, something constructed from many sources and stitched together in a not-fully-integrated way. More than anything else, Moenssens' article seems to be a collection of quotable "sound-bites" that may not fit together very coherently, but that can be dropped into briefs as quotes to provide the appearance of academic authority in response to any in-court attack on the dependability of handwriting identification expertise. The structure of Moenssens' article means that, in fashioning a response, one cannot merely rely on answering large points, but must take on particulars virtually paragraph by paragraph, lest the failure to respond be unfairly cited in the future as a concession of the validity of some isolated paragraph or other.[165] It seems clear that the intended audience of Moenssens' piece is not the academic community, whose members will generally read the piece in full if it bears on their work, and thus be in a position to evaluate its worth, but the judiciary, whose members (unfortunately but predictably) will not.

With this in mind, Part II will proceed as follows. Section A will deal globally with three of Moenssens' general attacks, because these attacks demand a global response. In section B we will then make our way tediously through Moenssens' article, responding in detail as we go. It is expected that Part II will serve more as a resource than as a source of pleasure reading, though there may be nuggets of amusement concealed therein to reward the casual reader.

---

[164] Many specific inconsistencies and infelicities are dealt with *infra*. For now, consider the following statements, which occur within two paragraphs of each other: "the task of comparing handwritings constitutes but a small part of the varied workload of questioned document examiners." Moenssens, *Post*-Daubert *World*, *supra* note 1, at 258. "Clearly, however, handwriting examinations occupy the major portion of many questioned document examiners' workloads." *Id.* at 259. Or contrast the assertion that "the progression of knowledge in handwriting examination has been phenomenal during the past twenty years" on pages 317-18, with the listing of the leading treatises in the field provided in notes 15 and 21, which bear the publication dates 1926, 1929, 1956, 1964, 1966, and 1982.

[165] We here declare that we are not confident that we have addressed every detail raised in Moenssens' article, and do not intend accidental silence to be any admission whatsoever.

A. Bias, Money, and Lack of Qualifications

Moenssens seems to claim that the *Exorcism* article can be viewed as part of a scheme on our part to make money by becoming expert witnesses even though we possess no expertise, and that even if our motivation might have been sincere in 1988, our interest in continuing to make money prevents us from seeing the light shed by subsequent research. He further claims that we have concealed our true motives and biases in our academic writings. Moenssens hammers away at these themes throughout the article, often supporting them with supposedly factual assertions that are, as we shall see, just plain wrong.

Once again, responding is made difficult by the somewhat unstructured and elusive nature of Moenssens' attack. To demonstrate this for the benefit of the reader unfamiliar with the Moenssens article, we have collected his assertions on these themes, and so as not to burden the main text with too much repetition, we have set them out in the extensive footnote below.[166] Our response follows. We believe

---

[166] Moenssens starts the theme of bias/money/lack of qualifications in the introduction to his article by saying:

> Since the publication of *Exorcism*, some of these authors have also become advocates and "expert witnesses" on the inexactness of handwriting identification. Portions of their testimony in several trial courts will be quoted herein, since the authors' answers on cross-examination provide us with a certain degree of "peer review" which was lacking at the time of the publication of their 1989 article. None of the authors has received training or acquired experience in handwriting identification techniques; they limited their inquiries to study of some of the handwriting identification literature. Their only "qualifications" to testify as "experts" on handwriting identification rest on the literature study and writing the 1989 article. In their more recent writing, the authors fail to reveal their bias to act as advocates in criminal litigation and oppose government document experts .... While no court has, as yet, excluded the testimony of qualified handwriting experts ... a few courts have, on the strength of the law professor-authors' testimony as expert witnesses, at least accepted their argument that handwriting identification is non-scientific.

Moenssens, *Post*-Daubert *World*, *supra* note 1, at 254-55 (footnotes omitted). Footnote 7 goes on:

> The two articles (and the book chapter) present some valid criticisms and make some worthwhile suggestions for improvement, while at the same time acceptance of their constructive criticism has become more difficult in light of the partisan positions in litigation taken by the authors in what appears to be a full-scale attack (misdirected in the opinion of this author) against most of the traditional crime laboratory expert functions.

Footnote 11 expands on the claim that Professors Saks and Denbeaux lack qualifications to testify in opposition to assertions of validity by questioned document examiners and expresses obvious displeasure that "'Professor Denbeaux was still testifying, attacking the reliability of handwriting identification evidence in trials as late as June, 1997 . . . .'" In footnote 36, Moenssens renews his assertions that Denbeaux should not be found qualified to testify concerning handwriting identification reliability, quoting James Starrs as saying that Denbeaux is "a lawyer who was out moonlighting as a document examiner with an honesty chip on his shoulder. His qualifications for his task as an expert were just a shade above that [sic] of a well read person." *Id.* at 263 n.36 (quoting James E. Starrs, *Recent Developments in Federal and State Rules Pertaining to Medical and Scientific Expert Testimony*, 34 DUQ. L. REV. 813, 830 (1996)). In footnote 43, Moenssens refers to Denbeaux as "the testifying law professor," while charging that his testimony in a case did not display an understanding of the differences between various schools of handwriting identification. That note is curiously repeated virtually verbatim as part of footnote 87. On page 274, Moenssens claims that the public silence of QDEs regarding *Exorcism* was a tactic that was a miscalculation "in light of the article's authors' move from the perch of academic inquiry into partisan advocacy." Footnote 111, and accompanying text, claims that "some" of the authors of *Exorcism* have asserted in expert testimony that only testimony that can be shown to be reliable by reference to calculated error rates and statistics should be admitted. Footnote 216 refers to the *Starzecpyzel* and *Velasquez* cases as cases "in which one or more of the *Exorcism* authors testified or sought to testify as expert witnesses against the admissibility of the handwriting experts' testimony." On page 305 Moenssens writes that the "debate about *Daubert*'s meaning in the context of 'scientific evidence' and 'forensic science' also offered the *Exorcism* authors an opportunity to descend from the Mount Olympus of academic anonymity into the public battle of the courtroom to advance their idea that handwriting identification was non-scientific . . . ." Footnote 235 says:

> At least one of the *Exorcism* authors, Mark P. Denbeaux, had according to his resume introduced in evidence on the occasion of his appearance as an 'expert witness' previous courtroom experience as a legal aid and pro-bono advocate for criminal defendants. Indeed, the genesis of the *Exorcism* article seems to have been a brief in the Mayflower Madam case, *People v. Barrows*, which Professor Denbeaux tried with research performed by *Exorcism* co-author Dr. Saks.

Footnote 237 is yet another extended but repetitive attack on the "authors" expert qualifications, as shown by "transcripts of Professor Denbeaux's testimony in several trials." Footnote 238 asserts that "despite their often professed public statements that their expertise extended only to the 'non-scientific nature' of handwriting comparisons, and that they did not profess to be handwriting examiners, one of the *Exorcism* co-authors (Professor Mark Denbeaux) testified to the invalidity of a specific identification." Footnote 248 states:

> The "Osbornians" among the questioned document examiners, under attack for what they perceive to be unjust and unfair criticism of their profession since the 1989 *Exorcism* article, have mounted a counter-attack against Risinger and his co-authors after the professors left the lofty spires of academe to descend into the pit of courtroom battle. Having become litigation antagonists, testifying as expert witnesses on the non-scientific nature of handwriting identification testimony for the defense in criminal cases against state and federal crime lab experts who take a differing view, it is only natural for Risinger and his co-authors to use any means to buttress the positions they take in litigation and direct their own ire at the "orthodox Osbornians" while giving credit to the graphoanalysts.

On page 308 Moenssens writes:

> What is particularly disturbing is that in the 1997 law review article, they also fail to reveal the clear bias they have shown by offering their

> services as expert witnesses "to attack handwriting identification" [quotes in original without further reference] any place in the country where their services are in demand. The intentional concealment of such bias for the sake of creating the appearance of scholarly equanimity and objectivity is indeed a serious defect that is antithetical to credible scholarship.

Footnote 251 continues:

> Thus, they fail to reveal in *Science and Nonscience*... that they were the experts who appeared or sought to appear as experts in the *Starzecpyzel, Velasquez* and *Ruth* cases... even though they cite the holdings in the cases with professorial approval. They also fail to mention that former co-author Mark P. Denbeaux, a colleague listed as being on the same faculty as Professor Risinger, continues to this day on the "have-expertise-will-travel" circuit-for-pay in criminal cases on behalf of defendants. He charges $250 per hour for his services in attacking the reliability of handwriting identification, or at least $3000 in one particular case.

On page 309 Moenssens writes "[t]he authors lacked the necessary understanding of the professions they sought to analyze, they lacked the objectivity that one expects of serious researchers, and they lacked the 'expertise' to evaluate competently the 'proper' methodologies for questioned document examiners or any other forensic scientist." Footnote 260 quotes a speech to a document examiner's group by attorney Thomas Black, (author of BLACK'S TEXAS EVIDENCE MANUAL) where he stated:

> The [*Exorcism*] authors do not appear to understand the process that document examiners go through to reach the conclusion that a signature is or is not authentic. The authors assume that lay people and document examiners use the same process to reach a decision. It is this erroneous assumption that makes their article as inexpert as their statement that document examiners are inaccurate. When they state that "no court has ever explicitly considered the field of document examination" one must respond, "neither did the authors."

Footnote 261 again attacks Denbeaux's qualifications, again setting out the Starrs quote *supra*. Footnote 280, in commenting on the methodological criticism of Kam I in *Science and Nonscience*, says:

> Stating the above was clearly prelude to the inevitable denigration, by a non-scientist lawyer the work of a true scientist [sic]. The pretense of offering *Science and Nonscience* as a scholarly "expose" of handwriting examination evidence, rather than a thinly disguised 'brief' in favor of their own advocacy positions, is more offensive since the fact that the authors use the articles as providing them with 'expert qualifications' in the broad handwriting field is carefully concealed from the reader. Thus, in citing the *Starzecpyzel* case, no mention is made that it was Professor Saks, co-author of both articles, who was the main "expert" witness, nor was any mention made anywhere in the article that Professor Denbeaux, co-author of *Exorcism*, continues traveling around the country and testifying in every case in which the defendant can pay the $2500 minimum fee that he charges per case. In several cases, Professor Denbeaux testified that he charged $250 per hour, but that in indigent defendant cases there was a $2500 limit to the fee he could charge.

Footnote 281 says:

> When a person adopts a categorical stand denying the legitimacy of a profession, as the *Exorcism* authors have done, and further takes it to the trial trail to advocate that position in litigation over a period of

we have been reasonably thorough in collecting the scattered threads of the bias/money/lack of qualifications attack in the footnote. We trust that the reader will also begin to see some of the problems of framing a response. Moenssens' attacks are certainly emphatic, but are neither organized nor entirely consistent. Is it Moenssens' position, for instance, that all three of us have testified as expert witnesses,[167] or is it that only some of us have done so, as other passages seem to say? The fact is, in the slightly less than 10 years since the *Exorcism* article appeared, Denbeaux has testified in 15 cases, Saks in one case, and Risinger in none (though he did consult with the defense in the *Starzecpyzel* case).[168]

We fail to comprehend the basis for Moenssens' general allegations regarding our "concealment" of relevant information in any of our writings.[169] The positions we took in *Exorcism* were clear, unambiguous, and public. If a litigant thought it would be helpful to have those positions explained in court so that better informed decisions might be reached, how is it a criticism of our position that one or the other of us agreed to do so? And as far as money is concerned, it seems little criticism that the person testifying should be compensated for his time.

Besides getting paid, our greatest sin, individually or collectively, seems to have something to do with the following words: "advocate," "bias," "partisan," "non-objective," and "pro-defense."[170] In reality, this sin reduces to our disagreement with Professor Moenssens. It is, without doubt, true that we believe some things, as each of our prior articles and Part I of this article have been directed toward elucidating. Since we believe the things we say, hedged with qualifications though they may be, we are not reticent to explain them in such a

---

> years, there is an understandable (though unjustified) desire to keep rejecting, systematically, the efforts of anyone engaged in presenting evidence showing that the advocate's premises were false.

Finally, footnote 347 in the conclusion states:

> Their opinions must be dismissed as unwarranted, untenable, and erroneous. Perhaps the critics are also motivated by the strong pro defense/anti-prosecution bias which they carefully concealed from the readers of their "articles." Some also engage in a profitable sideline of appearing as expert witnesses, relying on these purportedly carefully researched "studies" as their badge of expertise.

[167] *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 308 n.251.

[168] What may really be motivating Moenssens is not that two of us have testified, but that the testimony has had some effect. For instance, in nearly half of the cases in which Denbeaux has been involved, either the expert has been withdrawn or the result has been inconsistent with the expert's position.

[169] *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 308.

[170] *See, e.g.*, *id.* at 314 n. 281, 330 n.347.

way so as to move others to come to share our view by reason and reflection. This seems to be what Moenssens damns as advocacy. We thought that it was simply the nature of reasoned discourse, in academia or in the courtroom. Because we believe what we believe, according to Moenssens, we are "biased." The invocation of the word "bias" conjures up the related notion of "prejudice." However, we like to think of our positions as "postjudices" — positions arrived at after reflectively weighing such evidence as is available. Moenssens charges our attitude with being "non-objective," especially as it relates to our treatment of the Kam studies, and charges us with holding on to our publicly espoused positions in the face of overwhelming new evidence.[171] We will leave the evaluation of Moenssens' claims to the readers of Part I, and of *Science and Nonscience*, except to assert once again that we have attempted to be as skeptical as the needs of science and the legal system's stated ideology require while giving credit where credit is due, even to personally hostile authors such as the Galbraiths.[172]

Finally, Moenssens charges us with a "pro-defendant bias" in criminal cases. We believe we have a "pro-accuracy" bias within the context of the notion of proof beyond a reasonable doubt. We would be delighted to have well-validated sources of information that would dependably separate the innocent from the guilty, and subject the guilty to substantial punishment. We just don't believe in indiscriminately sacrificing defendants to maintain the appearance of such a system when evidence of the validity of that system's methods is lacking. If this has become Moenssens' definition of "pro-defendant," it marks a change of position for him,[173] and one which we view as unfortunate.

On the issue of our failure to "reveal" our biases,[174] in the general sense discussed above, we wear our positions on the sleeve of everything we write. In the specific sense, in *Exorcism* we revealed our previous litigation involvement dealing with handwriting issues,[175] and in *Science and Nonscience* we stated Denbeaux's identity as the expert in the *Velasquez* case in the text.[176] The principal author of *Science and Nonscience* (Risinger) has never testified as an expert, and the other author, (Saks) has testified only twice (once before and once

[171] *See id.* at 314.
[172] *See supra* note 25.
[173] *See supra* note 107.
[174] *See, e.g.*, Moenssens, *Post-*Daubert *World*, *supra* note 1, at 255.
[175] *See* Risinger et al., *Exorcism*, *supra* note 2, at 776-77.
[176] *See* Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 32.

after the publication of *Exorcism*). We do concede that it would have been better to have indicated in *Science and Nonscience* that Professor Saks testified in *Starzecpyzel*, if for no other reason than to have eliminated the single scrap of factually accurate complaint available to Moenssens. Frankly, it didn't occur to us as being an issue, since all we dealt with there in regard to the decision in *Starzecpyzel* was a textual analysis of Judge McKenna's written opinion.

However, we would like to point out the gross double standard Moenssens applies to us, as contrasted with Professor Kam. Kam has testified as much, or almost as much, as Denbeaux, a fact about which Moenssens is silent. Indeed, Kam and Denbeaux have often been called by opposing sides in the same case. Yet Kam has not disclosed any of this in any of his publications. In fact, in Kam II and Kam III he references the proceedings in eleven cases without indicating that he testified as an expert in ten of them.[177] Frankly, we never considered it much of an issue, but Moenssens' failure even to note it, much less to manifest the same outrage at it as he has at what turns out to be our single omission, speaks tellingly about what is really going on in his Article. As we have already noted, consistency and even-handedness are not Moenssens' strengths.[178]

While we are on the question of Moenssens' consistency, let us extend the analysis from failure to disclose to making money by being an expert[179] and "'qualifications' to testify as 'experts[.]'"[180]

Moenssens does seem genuinely outraged that Denbeaux would "oppose government document experts" and get paid for doing it.[181] At the same time, Moenssens totally ignores the fact that Kam has undoubtedly been paid for his testimony, without the fee limitations that apply to those testifying for indigent defendants. In addition, nobody paid us to write *Exorcism*, while all Kam's research has been

---

[177] *See* Kam II, *supra* note 8, at 785; Kam III, *supra* note 81, at 1003.

[178] Moenssens himself has served as an expert in regard to fingerprint issues at least as recently as 1994, and presumably he was paid for it. *See* Montana v. Cline, 909 P.2d 1171, 1177 (Mont. 1996). Yet, he does not explicitly note this in writings on that subject, such as the Fingerprint chapter in SCIENTIFIC EVIDENCE IN CIVIL AND CRIMINAL CASES, *supra* note 107, at 495-554. Moenssens has also represented a DNA laboratory, Lifecodes, in litigation, *see* Callahan v. Virginia, 379 S.E.2d 476 (Va. 1988), but has not revealed that fact in his writings on DNA dependability, such as Andre A. Moenssens, *DNA Evidence and Its Critics — How Valid Are the Challenges?*, 31 JURIMETRICS J. 81 (1990). Again, we don't think it is much of an issue, but it does seem, once again, that Moenssens has a severe case of double standards.

[179] *See, e.g.*, Moenssens, *Post*-Daubert *World*, *supra* note 1, at 255-56.

[180] *Id.* at 254.

[181] *See id.* at 255.

funded by the Federal Bureau of Investigation (FBI).[182] We want to make it clear that we are not criticizing Kam for this, but if either group is likely to have been influenced by monetary considerations, it is not the "Risinger-Denbeaux-Saks" group.[183]

At root base, Moenssens' arguments about our qualifications to evaluate evidence of validity are bankrupt. As we pointed out in Part I, Moenssens' hope is to disqualify pesky critics from testifying by requiring that only guild members be looked upon as competent to testify about the weaknesses of handwriting identification. Being a practitioner of handwriting analysis, however, has little to do with either designing tests to measure the performance of guild members according to a black box proficiency model,[184] or testifying to the implications of the weakness or absence of such testing. What is necessary for the above tasks is the effort to design and administer the tests or to gather the literature for review, and the critical skills to analyze and interpret the data. Compare the "Kam group" with the "Risinger group." Both have three working members. Both have one Ph.D. in science, Kam and Saks. Kam's degree is in Electrical Engineering and Computer Science, Saks's in Social Psychology (with an emphasis on research methodology).[185] No member of either group is trained as a document examiner or has any hands-on experience as such. Yet Moenssens views Kam not only as qualified to testify to the results of his research, but as a virtual demi-god, while Saks and Denbeaux are

---

182 Professor Kam's 1997 curriculum vitae reveals that his group had by then received $280,831 in grants from the FBI for handwriting related projects.

183 All QDEs testifying in support of their own expertise (by far the most common source of such testimony) are paid for their time, and earn their livings as a result of their claimed expertise. Elsewhere Moenssens himself has recognized the magnitude of this problem: "More and more, the courts are coming to recognize that . . . the foundation for admissibility and the fact of general scientific recognition 'may not be established without the testimony of 'disinterested and impartial experts,' 'disinterested scientists whose livelihood was not intimately connected with the new technique.'" Moenssens, *Novel Scientific Evidence*, *supra* note 115, at 5 (citations omitted). If personal monetary interest is the motivator that Moenssens claims it is in regard to us, it is many times that in regard to every QDE.

184 *See supra* note 107. Note that we are not saying that practitioners cannot or should not have input into test design. Practitioner input can ensure that tests and test materials are reasonably realistic and cover the range of problems most often encountered in practice. However, it is not a sine qua non, and much of what the researcher needs to know about such subtasks can be gained from reading the standard QDE literature. Kam had no apparent direct practitioner input in the design of his tests, at least none that is revealed in the text of his articles, and Moenssens has never criticized him for that.

185 On this score, Saks's degree is more pertinent, because the testing of QDEs deals with a human measurement problem. Kam's unfamiliarity with some of the special problems of research on human subjects may explain some of the design flaws in the Kam studies. *See supra* note 64 and accompanying text.

viewed in the most contemptuous terms, unqualified to testify to the results of their research because of their lack of training in document examination. To us, there seems to be little doubt that all of these persons should testify — better to inform the court and the fact finder of the data and its limits. But once again, Moenssens would apply inconsistent standards in order to silence critics and protect friends.

B. Faulty Testimony, Sloppy Scholarship and Other Delicts.

We think it is safe to say that Moenssens has accused the *Exorcism* article, and to a lesser extent *Science and Nonscience*, of bad scholarship using just about every negative adjective and rhetorical device known to humankind, including damning our honesty with faint denial.[186] But when we look at the content of the specific charges made, the charges begin to evaporate. Once again, as in the case of the bias/money/qualifications argument, Moenssens' position is scattered, not fully consistent, and hard to pin down. We will therefore attempt to collect all assertions about the deficiencies of our articles that are not merely name calling, but which have some verifiable content. We will begin with Moenssens' allegations asserting faults in our testimony in judicial proceedings, and continue from there. Unlike the approach we adopted regarding Moenssens' bias/money/qualifications argument, however, we will respond to each allegation in turn:

1. In footnote 8, Moenssens quotes a passage from Denbeaux's testimony under cross-examination in *United States v. Martin*[187] as an admission that the *Exorcism* article reached "extreme conclusions."[188] Taken in context, however, Denbeaux's conclusions were simply that, as of the date of publication of *Exorcism*,

> I always took two positions: either their [the FSF] tests are invalid and worthless or they are valid. The view turns out to be one of two things: There are either no studies defending the ability of handwriting experts to do that which they claimed they could do,

---

[186] Moenssens' footnote 293 begins, "Lest one accuse them of intentional dishonesty, there were perhaps a number of understandable reasons for their confusion . . . ." *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 317 n.293. It is perhaps worth noting here that Professor Moenssens apparently holds no college or university degrees except his J.D. from Chicago Kent Law School and his LL.M. from Northwestern University.

[187] No. 1:96-CR-287-JEC (N.D. Ga. Jan. 22, 1997); *see Martin* Transcript, *supra* note 16, at 295.

[188] *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 255 n.8.

or the studies that do exist, their studies seem to show they can't.[189]

That is, Denbeaux testified that the conclusions were extreme because the facts were extreme, not because the conclusions were unwarranted. Moenssens also quotes Denbeaux as saying that, in light of Kam II, those conclusions "require readjustment."[190] Certainly it is true that the "no data to negative data"[191] conclusion of *Exorcism* must be readjusted in light of Kam II, but as Denbeaux said immediately thereafter in *Martin*, it is unclear that Kam II is "going to lead to the change that the handwriting experts want . . . ."[192]

2. Later, in footnote 242, Moenssens tries to link the "extreme conclusions" language just discussed to a passage much later in the same testimony, where Denbeaux, being cross-examined on unpublished data he had received only the week before (which was later to become Kam II), said that the conclusion of *Exorcism*

> was an accurate depiction of the records at the time I wrote it, and by no means am I saying it was inaccurate then. There is

---

[189] *Martin* Transcript, *supra* note 16, at 294.

[190] *See* Moenssens, *Post*-Daubert *World*, *supra* note 1, at 255 n.8.

[191] *See* Risinger et al., *Exorcism*, *supra* note 2, at 750.

[192] The full context of the "extreme conclusions" quote, clearly referring to the quoted conclusions in the text from page 294 of the transcript, occurs on pages 295-96:

> Q. Do you consider your law review article to be scientific?
> A. Well, no, I wouldn't think it was scientific. It's a law review article. It's an analysis of data. I think it's rational and logical and well-founded and well-supported and its conclusions certainly are documented, but it is not science. I am a law professor studying the law that exists and the data that exists and I report it.
> Q. The data that existed at the time that you wrote the article?
> A. Sure.
> Q. Based upon the testimony that you have heard today, as well as information that you heard several weeks ago in a hearing here, U.S. versus Humphery, is it your understanding now that your law review article is outdated?
> A. Well, certainly it's outdated. I don't know that that means that its value is gone, but it's six years out of date if that's what you mean.
> Q. Are your findings accurate?
> A. Oh, the more data that comes in, it seems to me the more that the extreme conclusions of our law review article probably require readjustment. I would say that Dr. Kam's last study certainly requires some change. I am not sure it's going to lead to the change the handwriting experts want, but it will certainly change the differences in the numbers.
> Q. Is it not true now that there have been empirical studies done or empirical studies done on the proficiency of document examiners?
> A. Yes . . . .

Martin Transcript, *supra* note 16, at 295-96.

> more data. And the data obviously changes those numbers and it clearly is not nearly as hideous as the situation in which handwriting examiners were virtually no better than flipping a coin.[193]

The trouble with Moenssens' argument is that the "extreme conclusions" language did not refer to this, and that *Exorcism* reached no such "extreme conclusions."[194] *Exorcism* did show that, under one aggregation strategy, document examiners were right only fifty-seven percent of the time,[195] and also, for certain tasks in the FSF studies, document examiners did no better than guessing. This conclusion was confirmed by the Galbraiths's statistical analysis, which established that document examiners did no better than chance on two of six subtasks reflected in the FSF studies.[196] As we have explained in Part I, Kam II had no direct bearing on those subtasks.[197] To that extent, Denbeaux bent over backward in his quoted statement to be cautious and generous in the face of newly presented, unpublished data.

3. In footnote eleven Moenssens says, noting the decision of the Third Circuit in *United States v. Velasquez*,[198] that this case ruled that "it was an error to refuse to permit the defendant to call professor Mark P. Denbeaux as an expert 'critic' of the field of handwriting analysis, *and as a handwriting analyst.*"[199] On the contrary, the Court made no such ruling regarding Professor Denbeaux's competence as a handwriting analyst, as footnote 6 of the opinion explicitly states.[200] Rather, Denbeaux was found to be an expert on the "limitations of handwriting analysis."[201] Moenssens then says,

> The decision in *Velasquez* to find error in excluding Denbeaux's testimony was, at least in part, based on the Hon. Judge Roth's mistaken belief, *from his testimony*, that he was a social scientist and statistician, which he was not. Denbeaux, *when apprised of the judge's erroneous factual assumptions, failed to notify her of the discrepancy.*[202]

---

[193] *Id.* at 313.

[194] *See supra* note 55 and accompanying text.

[195] *See* Risinger et al., *Exorcism, supra* note 2, at 748.

[196] *See* Galbraith et al., *supra* note 25, at 15 tbl.3.

[197] *See supra*, text accompanying note 54.

[198] 64 F.3d 844 (3d Cir. 1995).

[199] Moenssens, *Post*-Daubert *World, supra* note 1, at 255 n.11 (emphasis added).

[200] *See Velasquez*, 64 F.3d at 848 n.6. The court stated that: "[W]e need not address the issue of whether Professor Denbeaux was qualified to testify as to his ability — or inability — to identify the handwriting exemplars proffered by the Government." *Id.*

[201] *See id.* at 846.

[202] Moenssens, *Post*-Daubert *World, supra* note 1, at 255 n.11 (emphasis added).

It is true that Judge Roth's opinion in *Velasquez* contained a footnote stating, among many other things, that Denbeaux had testified that he had "spent four years as a statistical social scientist."[203] In fact, what Denbeaux had testified to in the proceeding in the court below was that, while on the faculty at Seton Hall University School of Law, he had spent four years as an American Bar Foundation fellow working with co-author Dr. Alan Katz on a social science research project concerning the effects of law school on student cynicism, and that this project involved the gathering and statistical analysis of data. Partially as a result of significant health problems, Denbeaux paid little attention to the terms of the *Velasquez* opinion when it was issued, most particularly the footnotes. When he noticed the wording of the footnote in the fall of 1995 (he was not "apprised"), he consulted Charles Sullivan, Associate Dean of Seton Hall University School of Law, as to whether the court's statement was sufficiently inaccurate to require response. To be cautious, Denbeaux then called the Third Circuit clerk's office to inquire about the proper procedures to follow in order to address the issue, and was told that it was too late for anything to be done. Moenssens knew all of this, since it was explained in Denbeaux's testimony in the *Martin* case, which was cited elsewhere by Moenssens.

4. Later, in the same footnote, Moenssens says:

> Professor Denbeaux was still testifying, attacking the reliability of handwriting identification evidence in trials as late as June of 1997, ready to criticize on the witness stand a more recent empirical study by a respected university-affiliated research scientist, Dr. Moshe Kam, that supported the abilities of document examiners in identifying handwritings of individuals, *even though Denbeaux admitted only having heard about the study an hour earlier when Dr. Kam testified and without having studied the Kam report itself.*[204]

The highlighted statement is a total fiction. Perhaps it resulted from an unaccountable conflation of the details of three different cases, *United States v. Humphery*,[205] *United States v. Martin*,[206] and *United States v. Paul*.[207] The "more recent study" Moenssens refers to must be Kam II, since Kam I had been published and Kam III had not yet been produced. Denbeaux first heard about Kam II not through Kam's testimony, but when he was confronted with it on cross examination

---

[203] *Velasquez*, 64 F.3d at 847 n.4.

[204] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 255 n.11 (emphasis added).

[205] No. 94-CR-447-JEC (N.D. Ga. 1997).

[206] No. 96-CR-287-JEC (N.D. Ga. 1997).

[207] No. 97-9302 (N.D. Ga. 1997).

in *Humphery*, even though it had not been published or previously produced in discovery, and at that time Denbeaux refused to make any substantive comment at all until he was given an opportunity to read it.[208] Denbeaux was supplied with a copy of Kam II a week later (the prosecutor would not let him keep the one he was shown in *Humphery*), and when he testified concerning certain weaknesses that his preliminary evaluation revealed, such as the incentive problem discussed in Part I, *supra*, he had possessed the report for almost a week. (Kam later admitted that those criticisms, and related criticisms by Saks, were the reasons for Kam III.)[209] It is true that Denbeaux further criticized Kam II in *Paul*, in June of 1997, but by that time he had possessed Kam II for nearly four months. Thus, Denbeaux did not first hear of Kam II from Kam's testimony, and never testified critically concerning anything he had only seen an hour before, or that he had never studied. How Moenssens could innocently have confused the record to this degree is for the reader to evaluate.

5. In footnote 221, Moenssens states that, despite lack of qualification, Professor Denbeaux "proffered an opinion as to the invalidity of a specific identification in one case."[210] If by this he means that Denbeaux testified about the weaknesses of a QDE's identification and the dangers of relying on it, it is true that Denbeaux has done that in a number of cases, as *Velasquez* held he was qualified to do. If Moenssens means to say that Denbeaux offered his affirmative opinion that the defendant in a given case *did not* write the questioned document, this has never happened, and it never will happen.

6. Returning to footnote 11, Moenssens says of Saks's testimony in *Starzecpyzel*, that he "admitted that he had not read all of the technical literature"[211] and further that "[h]is testimony also showed that he was unaware of the terminology in use in the questioned document examination profession."[212] The "technical literature" referred to consisted of various writings by handwriting identification experts of a non-empirical nature, and the terminology was the specialized terminology reflected in that literature. Professor Saks was not testifying as a document examiner, but rather as a social scientist and research methodologist conversant with human skills testing issues. As

---

[208] Transcript of Proceedings at 32, 93, 94, United States v. Humphery, No. 94-CR-447-JEC (N.D. Ga. Jan. 8, 1997).

[209] *See Gilreath* depositions, *supra* note 64, at 48-49 (deposition of Kam).

[210] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 300 n.221, (citing Transcript of Proceedings, United States v. Pravato at 1598, No. 95-CR.981 (E.D.N.Y. 1996)).

[211] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 255 n.11.

[212] *Id.*

such, the things that Saks "admitted" he had not read were as irrelevant to his witness function as *Finnegans Wake* (which he also admits he has not read).

7. In addition, Moenssens asserts in footnote 11 that Saks "conceded that in formulating his *Exorcism* findings, *he had made a selective use of 'short answers' given to certain tests but had ignored the reasoning expressed by the persons tested*."[213] First, Saks never testified that he made "selective" use of anything. He did testify that in formulating the analysis in *Exorcism* we utilized the response categories assigned both by the test subjects and the FSF, and did not, based on narrative explanations, reassign answers to categories not assigned by the experts themselves.[214] We have dealt with this issue extensively in *Science and Nonscience*, and in Part I.[215] We feel that this decision was the right one, but whatever one's opinion on that, the decision hardly constitutes "selective use."

8. Referring to the fact that a version of *Science and Nonscience* appeared as a chapter in *Modern Scientific Evidence: The Law and Science of Expert Testimony*,[216] Moenssens suggests that there is something wrong or unusual in publishing a version of an article as a book chapter (referring to it as "recycling" in footnotes 6 and 249, and saying in footnote 215, "In an age that applauds 'recycling' as 'politically correct' such multiple use of one's literary (if not scholarly) product is no doubt a laudable civic pursuit"[217] (and he attacks *us* for sarcasm)). However, the book chapter was clearly identified as a version of an article previously published in the *Iowa Law Review*,[218] and this is neither uncommon nor dishonorable. For instance, in the same treatise, Kaye and Freedman's chapter on statistical proof was a revised version of a chapter in the Federal Judicial Center's Reference Manual on Scientific Evidence.[219]

9. Moenssens states (without reference to any specific language) that in our "more recent" writings (presumably *Science and Non-*

[213] *Id.* (emphasis added).

[214] This appears to be what Moenssens means by "short answers," though these would best be characterized simply as the "answers" to the questions that were asked, the narrative being merely supplemental. Where Moenssens gets the phrase "short answers," which he sets out in quotation marks, is a mystery.

[215] *See supra* note 57.

[216] *See* MODERN SCIENTIFIC EVIDENCE, *supra* note 40, § 22.

[217] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 299 n.215.

[218] *See* MODERN SCIENTIFIC EVIDENCE, *supra* note 40, § 22 at 79.

[219] *See id.* § 3 (stating in a footnote that the chapter "is an expanded and revised version of David H. Kaye & David A. Freedman, *Reference Guide on Statistics*, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 331 (Federal Judicial Center ed., 1994).").

science), we "argue that because handwriting identification is 'non-scientific' according to the *Daubert* model, it ought to be barred from the courtroom and that such a conclusion legally flows from recent case law."[220] This is an example of a mischaracterization undertaken in order to subject us to ridicule for positions we did not take. Nowhere have we made such a silly argument. What we do argue is that all expertise must be subject to the best criteria of validity of which the subject matter will admit, and that for forensic identification procedures like handwriting, if they cannot be validated internally as sciences they must be subject to external black box validity testing.[221] We argue that this has not been done sufficiently in regard to handwriting identification.

10. A related Moenssens tactic is to put things in quotes in such a way so as to suggest that we said them, when we never said such things. Thus, in footnote 222, Moenssens says, "The *Exorcism*. . .authors begin their 'destruction' of the handwriting identification lore by discussing the so-called 'Inbau Study' . . . ."[222] We never asserted that we were undertaking a "destruction" of the handwriting identification lore, or anything else. Similarly, on page 301, Moenssens writes, "While the *Exorcism* authors' methodology and thoroughness in research might be questioned seriously, what was wrong with their 'revelation' that handwriting identification did not represent "scientific knowledge?"[223] Again, we never claimed to be making a "revelation" about this, or anything else.[224]

---

[220] Moenssens, *Post-*Daubert *World*, *supra* note 1, at 255.

[221] *See supra* note 109.

[222] Moenssens, *Post-*Daubert *World*, *supra* note 1, at 300 n.222, 301.

[223] *Id.* at 301.

[224] Three other minor points may profitably be disposed of here without burdening the text. The first is the claim made by Moenssens in footnote 240 that at the 1996 Meeting of the American Academy of Forensic Science, we retaliated to audience hostility "in kind by referring depreciatingly and sarcastically to the questioned document examiners as 'you people better get your act together [sic].'" Those words were spoken, but by another panelist, not by any of us. The second is the assertion made in footnote 259 that "*Exorcism* began as a 'brief' in the trial of the Mayflower Madam." This is untrue. Denbeaux and Risinger first began reflecting on the validity of handwriting identification in the late 1970s in connection with an examination of the Lindbergh baby kidnapping case, and the issue was first raised in court as early as 1980, as was clearly indicated in *Exorcism*, *supra* note 2, at 776 n.196. The third point is the assertion in footnote 249 that we resist the truth revealed by the Kam research because it "removes the underpinnings of . . . [our] reasons for 'being' in the academic world." We would simply reply that each of us has other academic accomplishments that we like to think are at least as significant as whatever we have done in regard to handwriting identification.

11. In footnote 8, Moenssens begins his attack on the thoroughness of our literature search in *Exorcism*, which continues at various points throughout the article. This has been dealt with in Part I.A.1, *supra*.

12. In footnote 8, Moenssens further observes that "the article [*Exorcism*] was not peer-reviewed nor was it submitted to questioned document examiners or other scholars with experience in the forensic sciences . . ."[225] a theme that he repeats at footnote 232 where he says (while writing in a law review), "Need it be stressed at this time that law reviews are not peer-reviewed publications? The editors' work on a professor-authored article submitted for publication tends to be deferential — let alone one written by three law professors, especially if it is lengthy and full of extensive footnotes!"[226] How "deferential" the editors of the *University of Pennsylvania Law Review* tend to be is debatable, considering the fact that they reject significantly more than ninety percent of the articles submitted. However, this line of attack does raise an interesting issue, and one of the detail weaknesses of the *Daubert* opinion: the confusion of juried selection of articles for publication with the much broader concept of peer review. The norm in many academic areas is for journals to select articles for publication through juries of professionals or academics. This is the initial stage of "peer review." The most important peer review comes after publication as articles are digested, critiqued, and evaluated by the greater interested intellectual community.[227] The utilization of peer juries has strengths and weaknesses. It may often prevent publication of only the most transparently bankrupt research, since there are so many journals and such a variety of juries that almost everything apparently respectable will find a slot someplace. As to any given journal in the food chain of prestige, the jury system can be criticized for potentially institutionalizing single doctrinaire viewpoints.[228] Whatever the merits of such a system, law reviews are by tradition student juried, a system which has its own strengths and weaknesses. Many law reviews attempt to combine the virtues of both systems by involving faculty members in the evaluation process under various circumstances. In the case of the *Exorcism* article, it was reviewed and recommended prior to acceptance for publi-

---

[225] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 255 n.8.

[226] *Id.* at 304 n.232.

[227] At one point Moenssens suggests, perhaps tongue-in-cheek, that "answers on cross-examination provide us with a certain degree of 'peer review' . . .," Moenssens, *Post*-Daubert *World*, *supra* note 1, at 254, even though, of course, the purposes of objective peer evaluation and advocacy cross examination are radically different.

[228] *See* DANIEL MCNEIL & PAUL FREIBERGER, FUZZY LOGIC, 75-77 (1993).

cation by Stephen B. Burbank, David Berger Professor for the Administration of Justice at the University of Pennsylvania Law School, whom we are happy to count as a peer. The main point of Moenssens' attack seems to be that criticisms that cannot meet the approval of a jury of professional document examiners ought not to be published or paid any attention.[229]

13. On page 266, Moenssens asserts that "four unrelated occurrences" have "placed questioned document examination in the remarkable position of having to defend its place among the forensic sciences."[230] He then lists: "(1) the almost casual manner in which the law of expert handwriting evidence has developed in our courts from the earliest cases through modern times; (2) the growth of the substantial body of handwriting 'experts' whose professional roots originate in graphology and graphoanalysis"; (3), the *Daubert* decision; and (4), the publication of *Exorcism*.[231] In footnote 47 he then criticizes *Science and Nonscience* for crediting only two of the four, the *Daubert* case and the publication of *Exorcism*. However, it was *Exorcism* that first publicly documented the "casual manner in which the law of expert handwriting evidence has developed . . . ."[232] and in that sense the two are hardly unrelated or separate. As to the doctrinal rift between orthodox Osbornians and graphologically oriented practitioners, this rift was hardly known at all outside of document examiner circles before it was documented publicly in *Science and Nonscience*, so how it could have contributed to the existence of the legal controversy is unclear. We suppose Moenssens' argument is that all lack of data, all inaccurate results, and all poor performances are attributable to graphologists, so that is how any controversy developed. As we explained in Part I, however, the lack of data applies to all document examiners, and such negative data as exist come from tests administered largely to Osbornians.[233]

14. In footnote 87, Moenssens accuses us of "still" failing to "appreciate the width and breadth of handwriting identification 'types' and their differences in methodology and competency."[234]

---

[229] At another point, Moenssens admits that *Exorcism* has received generally favorable post-publication response, but claims that we merely pulled the wool over the eyes of such gullible scholars as Professors Margaret Berger, Ron Carlson, Paul Giannelli, Roger Park, Richard Rakos, and Stephan Landsman. *See* Moenssens, *Post-*Daubert *World*, *supra* note 1, at 298 n.208, 299 n.215.

[230] *Id.* at 266.

[231] *Id.*

[232] *Id.*

[233] *See supra* note 57.

[234] Moenssens, *Post-*Daubert *World*, *supra* note 1, at 273 n.87.

Presumably the "still" refers to both *Science and Nonscience* and to a quote from Denbeaux in the *Martin* case referring to "the range of handwriting experts who are out there testifying, some of whom may have done a great deal of work, others . . . very little . . . and most handwriting experts admit that there is a wide range of trained and prepared people performing as handwriting experts."[235] Denbeaux testified to no more than that which Moenssens himself repeatedly claims — that there is a wide range of training and preparation represented among those who testify in court as handwriting experts.[236] The major irony is that the first publication that generally outlined the main doctrinal split between Osbornians and graphology-influenced handwriting identification experts was *Science and Nonscience.*

15. On pages 297-98 Moenssens writes:

> Since finding and critically reviewing scientific material can be a daunting task for a lawyer, it is very tempting to accept statements in law review articles about scientific issues as gospel. Thus, little critical analysis of the accuracy of a law review author's factual statements and conclusions occurs. After all, nearly each statement in the law review article is justified by a footnote![237]

At this point Moenssens makes clear, were there any doubt, that he is referring to *Exorcism*, and continues:

> The practice of citing law review articles for scientific facts or technical conclusions is an undertaking fraught with the danger of unwitting error. The danger is magnified when the article's author is not a scientist or a specialist in a technical field, but either a lawyer or a law student writing a note or comment, purporting to evaluate the scientific or technical worth of a specialized discipline, rather than on the state of the law as it relates to that discipline. Thus, law review authors cite other law review authors who cite other law review authors . . . .
>
> Law professors also struggle with what constitutes scholarship . . . . When they write a lengthy article, such as the *Exorcism* publication, purporting to have done a thorough investigation of a technical field, they suddenly become the scholars on the subject. Thus it was the case [sic] with an article critiquing the valid-

---

[235] *Id.* (quoting *Martin* Transcript *supra* note 16, at 321).

[236] How is Denbeaux's testimony substantially different than this passage from Moenssens: "The task of comparing known and unknown handwriting specimens for the purpose of determining whether they were produced by the same author is currently practiced by a wide variety of individuals, of greatly differing backgrounds, training, and experience." *Id.* at 252.

[237] *Id.* at 297-98.

> ity of handwriting identification evidence. As might be expected, the article has been cited in other law reviews as 'gospel truth' for its conclusions, despite the limited acceptance it has received in the courts.[238]

This passage suggests that we were merely three random legal academics with no ability to examine technical science subjects at all. First, as Moenssens himself, somewhat inconsistently, is at pains to point out elsewhere in his article, academic credentials are not a sine qua non of critical abilities.[239] Second, neither Denbeaux nor Risinger will plead guilty to being so math and science phobic that they are incapable of making valid and instructive observations. However, the real kicker in this passage is how, as usual, Moenssens attempts to diminish or conceal the credentials and the role of Dr. Michael J. Saks in both *Exorcism* and *Science and Nonscience*. Saks is responsible for much of the technical text in both articles (and this one). In addition, he is responsible for evaluating and approving all text dealing with any scientific or statistical issues. He stands behind all of what has been said, and if Moenssens wants to get at Denbeaux and Risinger he will have to come through Saks. Given Saks's general qualifications,[240] the only remaining claim for Moenssens is that only a document examiner can offer any observations that ought to be

---

[238] *Id.* at 298-99 (footnotes omitted).

[239] "Thus, it is not the existence of academic degrees, or their absence, that determines whether scientific methods are being used." *Id.* at 321.

[240] Professor Saks earned his Ph.D. in social psychology, with an emphasis on research methodology and statistics. At the University of Iowa he is the Edward F. Howrey Distinguished Professor of Law and Professor of Psychology. He has regularly taught courses in research methodology and statistics, not only at Iowa but also at Boston College, Georgetown University Law Center, Ohio State University, Arizona State University, and the University of Virginia (where for a decade he taught the subject in an LL.M. program for appellate judges). Saks has had extensive experience reviewing empirical research studies in his role as editor-in-chief of a peer reviewed journal (*Law & Human Behavior*) and as an editorial board member and peer reviewer for numerous other journals. Similarly, he has evaluated research proposals for the National Science Foundation and regularly is called upon by various other organizations to serve as a reviewer of empirical social science research and of proposals for research. In addition, he has conducted numerous empirical research studies of his own and written literature reviews involving the evaluation and synthesis of empirical research studies on various topics, but mostly on studies of decision-making within the legal system. His empirical research findings have been relied on as authority in numerous judicial opinions, including several by the United States Supreme Court. He is the author or editor of 10 books and author of nearly 100 articles in both refereed journals and law reviews. The multi-volume work, MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY, *supra* note 40, which he co-edited and co-authored, has recently been characterized by evidence scholar John Strong as a work that "sets a new standard of excellence." John Strong, *A Guide to* Daubert, 82 JUDICATURE 39, 40 (1998) (book review).

given any credence, which of course is inconsistent with his near-canonization of Kam.

16. On page 300, Moenssens claims that the *Exorcism* article was published "under a titillating title that began with an allusion to witchcraft: the article *Exorcism*."[241] One does not exorcise witches, one exorcises demons in cases of possession. As the text of *Exorcism* makes abundantly clear, the title refers to the human tendency "under the stress of having to wrestle with important types of facts about which there is no good evidence" to create "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."[242]

17. On page 300, Moenssens further says that "the Inbau and Forensic Science Foundation studies that the authors reported on in *Exorcism* were admittedly defective ones," but then says "they were indeed not intended to be valid appraisals on [sic] the ability of handwriting experts to identify handwritings and it was improper for the *Exorcism* authors to cite them for that purpose."[243] The questions concerning the propriety of examining the FSF studies to see what, if anything, they might mean, is dealt with in Part I.[244] However, we will consider in more detail whether our treatment of the Inbau study in *Exorcism* was proper.

In footnote 222, Moenssens asserts that we were in error in referring to the Inbau study as a "'primitive and flawed validity study from nearly 50 years ago[.]'"[245] Moenssens then says that "[t]he gross mischaracterization of this study alone as a validity test of the abilities of document examiners casts serious doubt on the research methodology of the *Exorcism* authors."[246] He then reinforces this assertion with a quotation from a memorandum authored by Inbau five decades after the fact, that the Inbau study

> "was conducted merely to demonstrate the invalidity of the evidentiary rule at that time which permitted lay witnesses to express an opinion regarding a disputed signature or other piece of

[241] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 300.

[242] Risinger et al., *Exorcism*, *supra* note 2, at 782.

[243] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 300.

[244] *See supra* notes 54-55 and accompanying text.

[245] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 300 n.222 (quoting Risinger et al., *Exorcism*, *supra* note 2, at 738).

[246] *Id.*

handwriting based upon a memory comparison between it and the writing of a person which they had seen at an earlier time."[247]

Moenssens then concludes that this "was its only purpose; it was never intended as a validity study of anything, nor was it represented as such."[248]

Was it or was it not a "validity study?" The Inbau study involved testing how well various people could identify the genuineness of actual and forged signatures. The people were the "instruments" being tested. The types of people tested were law professors, their secretaries, bank employees, professional document examiners, and people of assorted other occupations. The criterion against which their conclusions on any given item were evaluated for correctness was the known authorship of the documents. This fits the description of what testing researchers would call a validity study. Perhaps this would have been more obvious if the study had included calculations of the correlation between the participants' conclusions and the criterion of ground truth. That correlation would be termed a validity coefficient in psychometric research, which in a very real sense is what the Inbau study is: measurement of certain abilities among people.

By its own terms, the Inbau study seems plainly to be a study of the relative accuracy (that is, validity) of different ways the law might try to determine the genuineness of writings. In the study's introduction, Inbau wrote:

> Although the present time is a rather late date in the history of our law of evidence to question the validity of so well established a practice as the identification of disputed handwriting by lay witnesses, the following discussion is devoted to an experiment conducted for the sole purpose of ascertaining the accuracy of this method of identification.[249]

Later in the article, Inbau explains:

> [T]he suggestion was made that it would be of interest to ascertain the degree of success [handwriting] experts could attain in their examination of the same signatures. With the scope of the experiment thus extended, the writer .... test[ed] ... professional experts ....[250]

---

[247] *Id.* (quoting Fred E. Inbau, Memorandum to all Faculty Members of Northwestern University School of Law (Aug. 23, 1989)).

[248] *Id.*

[249] The Inbau study, *supra* note 44, at 434.

[250] *Id.* at 436.

By including QDEs in his study, Inbau created the possibility of comparing the relative accuracy (validity) of lay people versus QDEs. Simply because the study was a poor one does not make it non-existent. Inbau's study did make a very limited comparison of QDEs using comparisons to non-experts and did evaluate the accuracy of all the people tested. If there is a more apt description of that aspect of the study (the aspect relevant to *Exorcism* and the present discussion) than that which we advanced by our having referred to it as a "primitive and flawed" "validity study," then by all means tell us, and we will be happy to use the more apt term.

Moreover, that the study's weaknesses for the purpose of our article led us to place no reliance on its findings would seem to render entirely moot Professor Moenssens' frenzied charges of "gross mischaracterization" and "misrepresentation." We gave the study zero weight. What less would Professor Moenssens have had us do?

18. Moenssens writes on page 301:

> *Exorcism*'s main premise that because the reliability of handwriting identification evidence could not be mathematically or scientifically established, a confidence in the ability of the experts engaging in handwriting analysis had not been verified statistically, and courts ought not to admit it, seemed a clear non sequitur. Handwriting evidence was already firmly ensconced in the courts, and such evidence could rely on a century and a half of judicial approval . . . .[251]

First, to the extent we understand what this sentence means, it does not represent what we said. What we said was, "If handwriting identification testimony were to be proffered and treated as a case of first impression now, the proponent would clearly have the burden of proving the existence of the claimed skill, a burden that has yet to be met in any forum — legal, scholarly or scientific."[252] We then went on to say, "Despite the undeniable power of this argument, we are not so naive as to think that courts will be receptive to demands for exclusion of such testimony . . . . They are likely to continue admitting such evidence just because it has always been admitted, at least within living memory."[253] We are at a loss to see how these observations qualify as non sequiturs. We might well respond with a Latin aphorism of our own: All we were arguing was "*ex nihilo nihil fit.*"[254]

---

[251] *Id.* at 301.

[252] Risinger et al., *Exorcism, supra* note 2, at 772.

[253] *Id.* at 773.

[254] "Out of nothing comes nothing," or more appropriately to the point of the *Exorcism* article, ought to come nothing.

19. In footnote 222, Moenssens attacks our decision to consider anything generated by the FSF studies in the following terms:

> The *Exorcism* authors' attack on the Forensic Sciences Foundation (FSF) studies, again, was totally unwarranted and evidence of poor scholarship. The FSF studies were, admittedly, poorly conceived, poorly administered, and poorly interpreted. However, to rely on the FSF studies in support of conclusions about the proficiency, or lack thereof, of document examiners as a profession is an absolute non sequitur.[255]

Again, our understanding of "non sequitur" and Moenssens' seems to be radically different, and the meaning of this passage in general is a little obscure. We did not "attack" the FSF studies, though we did note some weaknesses, which Moenssens appears to concede completely. Our main sin seems to be not attacking the FSF studies, but reporting them. However, Moenssens fails to report that we also said this:

> It must be noted here that FSF's Proficiency Advisory Committee disavows these tests as representative of the level of performance in any of the fields being tested. Because of the high level of anonymity maintained and the limited amount of information collected by the proficiency testing program, it is not known who takes the tests for any laboratory or what techniques they used. Accordingly, we use them only to answer one conservative question: Do these data provide any additional evidence of the existence of handwriting identification expertise?[256]

Also, as we noted in *Science and Nonscience*, if the FSF had regarded the tests as totally meaningless, they would have ceased to give them or to issue reports providing aggregate data.[257] One can be forgiven for inferring that the FSF disclaimer was at least in part politically motivated. They were attempting to obtain participation with their proficiency testing program, and the disclaimer attempting to prevent the use of the information for evaluative purposes on any greater level than the local lab level, at least by anyone outside the FSF, may have been seen as a stimulant to cooperation. The FSF's own proficiency committee reports treat the data as at least suggestive or provisionally meaningful, and our similar treatment after disclosure of the FSF's disclaimer is hardly "totally unwarranted and evidence of poor scholarship."[258]

---

255 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 300 n.222.

256 Risinger et al., *Exorcism*, *supra* note 2, at 744 n.47.

257 *See* Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 42.

258 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 300 n.222.

20. The specifics of footnote 222 introduce Moenssens' main and favorite rhetorical ploy — the McCarthyite claim that he knows of many errors of a type he illustrates with only one or two examples, which he asserts are typical, and which often turn out not to be errors at all. Regarding the treatment of the FSF studies in *Exorcism*, Moenssens writes:

> Even so, the *Exorcism* authors inaccurately reported on what had been published. See, e.g. the mischaracterization/misquote of the 1975 FSF test's results when stating that "4(5%) said they could not make any conclusion from what had been submitted." In fact, the actual text of J. Peterson . . . stated that "[f]our laboratories, or 5.4%, reported inconclusive results but specifically mentioned in their reports that they noted significant agreement between the questioned material and the exemplar handwriting of 'B.'"[259]

Moenssens then invokes the "junk scholarship" reference.[260] Please. We did not directly quote anything. We set out data in a data table. It mischaracterized nothing. Our table accurately reported four inconclusives, which is how these results were characterized both in the report and by the tested document examiners themselves. We presented accurate percentages, in whole numbers with no decimal places. This was the practice we followed in all of our data tables, because we deemed such decimals insignificant digits. We nowhere reported inconclusives as "leaners" based on test taker narratives, whether the test taker was leaning toward right or wrong answers. We simply accepted the bottom-line characterizations of the report and the submitting laboratory. We still believe this approach is the best one for the data, as we indicated in *Science and Nonscience*[261] (a fact nowhere noted by Moenssens). Indeed, it was the decision of the Galbraiths to go back and count inconclusives as correct answers in their article, a main point of debate between them and us, which we explored at length in *Science and Nonscience*.[262] While there may be debate about whether such responses should properly be counted as "qualified opinions" or merely "bet hedges," our approach was hardly evidence of "junk scholarship."

21. Moenssens criticizes us on page 301 as follows:

> The *Exorcism* authors, while studying so much published information, had missed more than they had examined since they con-

[259] *Id.* (citing Risinger et al., *Exorcism, supra* note 2, at 744) (citations omitted).
[260] *See* Moenssens, Post-*Daubert* World, *supra* note 1, at 330 n.348.
[261] *See* Risinger & Saks, *Science and Nonscience, supra* note 2, at 53.
[262] *See id.* at 50-58.

> fined their studies to the 'book' literature, which turned up graphological as well as true document identification texts. They missed most of the periodic literature and all of the voluminous private research papers about on-going research efforts that were being shared freely by professional document experts attending regular training sessions and association meetings around the country.[263]

In Part I, we dealt globally with the charge that our literature search was deficient. All that needs to be said here is that, as we noted in *Exorcism*,[264] we consulted Maureen Casey Owens, past president of the ASQDE and then chief document examiner for the Chicago Police Department, in formulating our literature search. We missed none of the periodical literature written in English, at any rate. And, as we continue to repeat, if there was any empirical data reported anywhere which we missed in 1988, no one has cited such a study yet, including Moenssens. His main criticism seems to be that we did not consult unpublished sources, for Moenssens goes on:

> [b]efore the fairly recent (1995) appearance of the specialized publication International Journal of Forensic Document Examiners, much of the research that was pursued by members of the ASQDE was circulated privately only within the perhaps less than 100 person group of highly skilled questioned document examiners.[265]

Moenssens continues:

> The relatively small size of the interested audience caused traditional journals to be less interested in publishing limited-interest research papers; therefore, the research remained known only to a small group of professionals. It was indeed in these training seminar papers and the modern articles and monographs circulated among the professionals that the bulk of the 'scientific' research was featured.[266]

Criticizing us for not discovering and examining such unpublished material borders on the ludicrous,[267] especially since Moenssens elsewhere in the article seems to claim that even published articles should be given short shrift if not selected for publication by peer ju-

---

[263] Moenssens, *Post-*Daubert *World*, *supra* note 1, at 301.

[264] *See* Risinger et al., *Exorcism*, *supra* note 2, at 783.

[265] Moenssens, *Post-*Daubert *World*, *supra* note 1, at 301 n.223.

[266] *Id.* at 301-02.

[267] Any implication that unpublished material would have been readily available must be evaluated in light of the recent siege mentality of the QDE community. The ASQDE website is passworded beyond the homepage, and Professor Saks has been refused copies of conference papers that he has requested.

ries. None of this, however, should obscure the fact that, as yet, no one, Moenssens included, has pointed out a single non-anecdotal empirical study that existed in 1988 that we missed.

22. In attempting to document his charge that *Exorcism*'s "tone was sarcastic" Moenssens cites the following:

a. Moenssens says, "Handwriting identification is referred to as a 'spot' within judicial decisions."[268] Actually, it is the judicial decisions themselves and their lack of analysis (or, as Moenssens himself characterizes them, "their almost casual manner")[269] that were referred to as forming a spot. The actual text reads:

> Most judicial thinking on how to deal with asserted expertise in the courtroom is a product of the 20th century. The law has not yet worked out a coherent theory of control, and the results of judicial decision have been spotty. The rest of this Article is about one of the spots.[270]

b. Most of the other examples that Moenssens proffers deal with our discussion of Inbau, Wigmore, and Osborn, and have been dealt with elsewhere.[271] One further fairly egregious mischaracterization is worth noting here, however. Moenssens says, "The authors stated that handwriting identification boasted 'no [] discipline . . . except in the most desultory, disorganized, nascent, casual jackleg [sic] fashion . . . .'"[272] The actual passage from which the quoted language was taken refers only to the period prior to the Common Law Procedure Act of 1854:[273] "As we have seen, however, no such discipline seems to have existed in English speaking countries prior to the passage of the statute except in the most desultory, disorganized, nascent, casual, jackleg fashion . . . ."[274] This is a statement even Albert Osborn would have agreed with, applied to its time and place. What, if anything, Moenssens sought to accomplish by the misleadingly edited and quoted passage, beyond further scandalizing us, is for the reader to determine.

---

[268] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 302 n.226 (citing Risinger et al., *Exorcism*, *supra* note 2, at 733).

[269] *See supra* Part II.6.

[270] Risinger et al., *Exorcism*, *supra* note 2, at 732-33.

[271] *See supra* text accompanying notes 151-62.

[272] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 302 n.226 (quoting Risinger et al., *Exorcism*, *supra* note 2, at 758) (alterations in Moenssens, *Post*-Daubert *World*).

[273] Common Law Procedure Act, 1854, 17 & 18 Vict., ch. 125 § 27 (Eng.).

[274] Risinger et al., *Exorcism*, *supra* note 2, at 758.

23. In attempting to document his assertion that we were "scornful" in our "speculative conclusory statements at times"[275] Moenssens cites the following, which are responded to seriatim:

a. Moenssens quotes us as saying, "'Because of complaints from document examiners that prior tests were too difficult, the Proficiency Advisory Committee decided to make the 1987 test easy.'"[276] We did say that. What Moenssens omits is the next sentence, quoting the FSF report: "According to the report, '[t]his test was designed to be a relatively easy and straightforward test, because of complaints about previous test design.'"[277] Exactly how were we here being "scornful" in our "speculative conclusory statements?" Where the scorn, and more importantly, where the speculation?

b. In Moenssens' second example of us being "scornful" in our "speculative conclusory statements" he quotes us as follows: "Inbau experienced considerable difficulty in obtaining the assistance of professional document examiners . . . . In certain instances the real reason was undoubtedly the examiner's unwillingness to subject his reputation to any sort of experimentation."[278] What Moenssens fails to note is that everything after the word "Inbau" *is a direct quote from Inbau's article, in Inbau's words, and is clearly indicated as such by us.*[279] Thus, if anyone was being "scornful" in a "speculative conclusory" statement, it was Inbau. We merely reported what he wrote. (We think he probably had good ground for his conclusion.) Again, what motivated Moenssens' erroneous attribution beyond the desire to scandalize is for the reader to decide.

c. Moenssens' last purported example of us being "scornful" in "speculative conclusory statements" is quoted: "'[T]here is at least some anecdotal historical evidence that prosecutors can negotiate with forensic scientists to turn inconclusive or even negative reports into something that sounds inculpatory in court.'"[280] We did say that in footnote 72 of *Exorcism,* but we then continued, "This is what apparently happened with the ballistician in the government's case

---

275 Moenssens, *Post*-Daubert *World, supra* note 1, at 302.

276 *Id.* at 302 n.227 (quoting Risinger et al., *Exorcism, supra* note 2, at 747).

277 Risinger et. al., *Exorcism, supra* note 2, at 747 (quoting COLLABORATIVE TESTING SERVICES, INC., CRIME LABORATORY TESTING PROGRAM, REP. NO. 87-5, QUESTIONED DOCUMENTS ANALYSIS 1 (1987)) (brackets in Risinger et al., *Exorcism*).

278 Moenssens, *supra* note 1, at 302 n.227 (quoting Risinger et al., *Exorcism, supra* note 2, at 741 n.44).

279 *See* Risinger et al., *Exorcism, supra* note 2, at 741 n.44 (quoting Inbau Study, *supra* note 44, at 440).

280 Moenssens, *Post*-Daubert *World, supra* note 1, at 302 n.227 (quoting Risinger et al., *Exorcism, supra* note 2, at 747 n.72).

against Sacco and Vanzetti," citing authority.[281] How these heavily qualified statements count as "speculative conclusory statements" is a mystery to us. Certainly the slanting of testimony in ways not justified by the data but favorable to the prosecution has been repeatedly documented, most recently by the Inspector General's report on the FBI Crime Laboratory,[282] and has been recognized as a danger by Moenssens himself in other writings.[283]

24. In documenting his assertion that we were "demeaning and depreciating,"[284] Moenssens cites the following examples:

a. He quotes us as saying, "'Like folk medicine, handwriting identification may sometimes be efficacious but no verification yet exists of when, if ever, it is and when it is not.'"[285] We did say that. In fact, we asserted that folk medicine was the best analogy for thinking about handwriting expertise given the state of the empirical record. We still think that. It is not "demeaning or depreciating," merely accurate.

---

[281] Risinger et al., *Exorcism*, *supra* note 2, at 747 n.72 (citing JOUGHIN & MORGAN, THE LEGACY OF SACCO & VANZETTI 15-16 (1976)).

[282] *See* U.S. DEPT. OF JUSTICE/OFF. OF THE INSPECTOR GENERAL, SPECIAL REPORT, THE FBI LABORATORY: AN INVESTIGATION INTO LABORATORY PRACTICES AND ALLEGED MISCONDUCT IN EXPLOSIVES-RELATED AND OTHER CASES (1997).

[283] Consider the following:

> [E]ven where crime laboratories do employ qualified scientists, these individuals may be so imbued with a pro-police bias that they are willing to circumvent true scientific investigation methods for the sake of "making their points" . . . . Unfortunately, this attitude is even more prevalent among some "technicians" (non-scientists) in the crime laboratories, for whom the presumption of innocence disappears as soon as police investigative methods focus on a likely suspect.

Moenssens, *Novel Scientific Evidence*, *supra* note 115, at 6.

And further,

> The temptation to fabricate or to exaggerate certainly exists. All experts are tempted, many times during their careers, to report positive results when their inquiries come up inconclusive, or indeed to report a negative result as positive when all other investigative leads seem to point to the same individual. Experts can feel secure in the belief that their indiscretions will probably never come to light.

Id. at 17. What would Moenssens have said about us if we had written these things?

Contrast these positions taken by Moenssens in 1993 with his assertions in *Post-*Daubert *World* that prosecution experts in general, and Osbornian handwriting experts in particular, always undertake comparisons either from a "result-neutral" position, or from an assumption "the opposite of what they are asked to determine." Moenssens, *Post-*Daubert *World*, *supra* note 1, at 322-23 nn.311, 312 and accompanying text. Also note his criticism of Denbeaux for testifying in terms that reflected, essentially, Moenssens' own 1993 position. *See id.* at 322. We believe that Moenssens' 1993 position in fact gave the more accurate picture of practical reality.

[284] *See* Moenssens, *Post-*Daubert *World*, *supra* note 1, at 302.

[285] *Id.* at 302 n.228 (quoting Risinger et al., *Exorcism*, *supra* note 2, at 734).

b. Moenssens further states "[i]n concluding, erroneously, that document examiners were correct 57% of the time and incorrect 43% of the time, the article states: 'This capriciousness contrasts with claims by document examiners that their field is just shy of perfect.'"[286] First, our conclusion was not erroneous, given the aggregation formula we used, which was fully explained in the text. The fact that there are potentially other ways to aggregate the data from five disparate tests does not render our conclusions "erroneous," and we still think ours was the best way to approach the data. Second, there is plenty of evidence that the position advanced by Osbornians for public consumption has been that the accuracy of handwriting identification, as they practice it, borders on perfect. In *Exorcism*, we cited Irby Todd advancing the above position.[287] In *Science and Nonscience* we cited the statement of FBI Document Section Chief Ronald Furgerson that all "180" "certified" document examiners in the United States would reach the same conclusions in any given case as he would, and we cited the trial testimony of FBI document examiner Richard Williams in *United States v. Smyth*, making a similar claim.[288] The available hard data refute those wishful claims, and that no doubt is one reason why Moenssens and the QDE community find these data so objectionable.

c. As another example of our being "demeaning and depreciating" Moenssens quotes us as saying, "'Listening to the evidence or reading the reports of experts, we often feel as if we were attending one of Mr. Home's seances. We are told that others see and we ought to see what we cannot see.'"[289] But we did not say this. This is a quotation from an 1871 letter to the London Times attributed to A. Hayward, Esq., Q.C., reviewing Charles Chabot's book *The Handwriting of Junius Professionally Investigated*, clearly identified as such by us.[290] This quotation was in the "history" section of the article.

---

[286] *Id.* (quoting Risinger et al., *Exorcism*, *supra* note 2, at 748 n.73).

[287] *See* Risinger et al., *Exorcism*, *supra* note 2, at 738 n.28 (citing Irby Todd, *Do Experts Frequently Disagree*, 18 J. FORENSIC SCI. 455, 458-59 (1973) (stating "it appears safe to answer a categorical 'No'" to the question of whether document examiners frequently disagree)).

[288] Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 41 n.100, 62 n.153 (citing DAVID FISHER, HARD EVIDENCE 196 (1995) and United States v. Smyth, 863 F. Supp. 1137 (N.D. Cal. 1994)).

[289] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 302 n.228 (quoting Risinger et al., *Exorcism*, *supra* note 2, at 760).

[290] *See* Risinger et al., *Exorcism*, *supra* note 2, at 759.

d. Finally, the fundamentally accurate and not disrespectful nature of our observation that Osborn's "mastery of penmanship offered a path off the farm" has been dealt with, *supra*, at note 159.

25. In documenting his assertion that *Exorcism* was "full of non sequiturs" Moenssens cites the following:

a. A non sequitur is "[c]ommenting that judges accord little weight to handwriting expert testimony because they always testify 'in favor of the party who called them,' even though witnesses 'of equal honesty, intelligence and experience' testified opposite to the experts."[291] Once again, Moenssens has attributed to us the words of somebody else. We have never said judges, in general, accord little weight to handwriting expert testimony. The point referred to by Moenssens (with the wrong page reference), and the quoted passages, are from the New York Court of Appeals decision in *Hoag v. Wright*.[292] The *Hoag* court was one of a number of courts quoted by us to illustrate "[t]he skepticism of some late 19th and early 20th century courts,"[293] a point for which the quoted language is certainly a proper illustration. We do not think it's a non sequitur either, but that is between Moenssens and the judges who decided *Hoag*.[294]

b. A non sequitur is describing "Osborn's departure from the farm to attend the Rochester Business Institute, his becoming a teacher of penmanship who realized he was 'in a dying profession' and decided to go into the handwriting expert business."[295] We have already voiced our puzzlement at Moenssens' notion of non sequitur, but we must note that, once again, he has his summary wrong. Osborn left his father's farm to join the faculty at the Rochester Busi-

---

[291] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 302 n.229 (citing Risinger et al., *Exorcism*, *supra* note 2, at 764).

[292] 66 N.E. 579 (N.Y. 1903).

[293] Risinger et al., *Exorcism*, *supra* note 2, at 762.

[294] This habit of attributing authorship to us of things we are merely quoting is irritating. We are beginning to suspect that Professor Moenssens relied too much on secondary sources, such as lists of complaints by others less academically experienced, rather than actually writing or checking these parts of the article himself. We have in mind the 1990 "four-page single-spaced list" by Charles C. Scott (which we have never seen), cataloguing, "serious errors and omission in legal citations" referred to by Moenssens in note 233, and "[c]ritiques of *Exorcism*" which were "shared within the questioned document examiners' professional gatherings and at their 'chest-beating' sessions." Moenssens, *Post*-Daubert *World*, *supra* note 1, at 302 n.225. We suspect that Moenssens relied on the above, together with the input from the "many forensic document examiners," some named and some not, who are credited with "unearthing information," and Peter V. Tytell, who furnished "unpublished materials." *See id.* at 251 n.**.

[295] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 302 n.229 (citing Risinger et al., *Exorcism*, *supra* note 2, at 765).

ness Institute, not to attend it. And as to the "dying profession" of penmanship, as we point out in *Exorcism*, even a person of less intelligence than Osborn would have figured this out pretty quickly after the introduction of the typewriter.[296] Training penman copyists in 1873 was like training typewriter repairmen in 1986: there was not going to be much of a future in it.

c. Another asserted non sequitur in *Exorcism* was drawn from an article published eight years after *Exorcism* (a nice trick in itself). Moenssens writes, "In *Science and Nonscience*. . . Risinger refers to Osborn's 1910 book, *Questioned Documents*, and suggests that its success was due to Osborn's friendship with evidence law giant John Henry Wigmore and the glowing review the book received by another legal giant, Roscoe Pound . . . ."[297] There is no assertion in *Science and Nonscience* that the success of Osborn's book was due to "friendship," and we don't really know how close Osborn and Wigmore were socially. It is clear, however, that Wigmore was a convert to Osborn's ideas, that he wrote an introduction to the book, and that he promoted it. It is hardly a non sequitur to suggest that Wigmore's support and the positive review by Pound contributed significantly to the book's acceptance and success, which is all that we suggested in *Science and Nonscience.*

d. Moenssens hauls out the "800 pound gorilla" comment here, which is dealt with at length *supra* in notes 152-55 and accompanying text (though it must be here noted that it is unclear how this comment qualifies as a non sequitur.)

e. Finally, Moenssens says:

> Also, in referring to the European origins of handwriting identification, it is remarked that it was from "[t]he same intellectual climate that gave us phrenology, Lombrosian physiognomy and, as previously noted, graphology, gave us 'chirography' or handwriting identification as a "science.'" The inference the authors must want the reader to draw is that because Europe spawned myths and fallacies, that everything that came from Europe must be of the same ilk.[298]

Now, if that is what we intended, that *would* be a non sequitur. However, the quoted passage was at the end of a discussion of the late 19th century continental attraction to "extreme rationalism on the

[296] *See* Risinger et al., *Exorcism*, *supra* note 2, at 765.

[297] Moenssens, *Post*-Daubert *World*, *supra* note 1, at 302 n.229 (citing Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 26).

[298] *Id.* (quoting Risinger et al., *Exorcism*, *supra* note 2, at 758-59) (internal citations omitted).

fringes of science, where theories were spun out to satisfyingly mystical complexity and experience was expected to conform or be damned."[299] It was in this context that the quotation given by Moenssens was written. We did not suggest that "Europe" had spawned these things, but a common mindset on the fringes of 19th century European science. So much for Moenssens' asserted "non sequiturs."

26. Next, in footnote 231 we encounter one of the most puzzling of Moenssens' positions, which is repeated later on page 307 and in footnote 248. In footnote 231 he says, "The authors of *Exorcism* proclaim that they did not discuss 'graphology' in their article . . . . However, they immediately thereafter proceeded to cite, in the very next note, a graphology book for a quote on the antiquity of the concept of individuality of handwriting."[300] Moenssens goes on, in the same note, to refer to the "egregious" "violation" of our "self imposed limitation on the discussion of graphologists"[301] because we noted in *Exorcism* the case of *Heller v. Murray*,[302] a case where a handwriting identification expert (who happened to be a graphologist, and incidentally a psychologist also, but both facts were irrelevant to the point for which we cited the case) refused to take a test during cross-examination that the judge had ruled appropriate.[303] Then on page 307, Moenssens says, "*Science and Nonscience* also continues the practice of on one hand, claiming to avoid dealing with or passing judgment on graphology, and on the other hand relying on statements by its practitioners to prove some point in Risinger and coauthors' criticism of handwriting identification."[304] Finally, Moenssens illustrates this in footnote 248 by pointing to our citation to Huntington Hartford for a quotation from Aristotle, our reference to Camillo Baldi's 1625 treatise (which we identified as of a "definite graphological cast") as the earliest modern European work asserting individuality of handwriting, our discussion of the animosity between "orthodox Osbornians" and graphologically influenced identification experts, our reference to a work by Robert Saudek for the proposi-

---

[299] Risinger et al., *Exorcism*, *supra* note 2, at 759 (footnotes omitted).

[300] Moenssens, *Post-*Daubert *World*, *supra* note 1, at 303 n.231 (citing Risinger et al., *Exorcism*, *supra* note 2, at 733 n.13 & 734 n.14).

[301] *Id.*

[302] 447 N.Y.S.2d 348 (N.Y. Civ. Ct. 1981). Moenssens also gleefully points out that *Heller* is "incorrectly cited." He fails to indicate that this "incorrect citation" is because the volume number of the secondary citation is given in *Exorcism* as 442 N.Y.S.2d instead of 447 N.Y.S.2d. *Nostra culpa.* We might here point out that Moenssens cited an article, Mark Davis, *Weird Science*, J. MARSHALL L. REV. 22 (1997), which is nonexistent. *See* Moenssens, *Post-*Daubert *World*, *supra* note 1, at 264 n.41.

[303] *See Heller*, 447 N.Y.S.2d at 350-51.

[304] Moenssens, *Post-*Daubert *World*, *supra* note 1, at 307 (footnotes omitted).

tion that graphologists believe their approach aids identification of handwriting,[305] and our notation of the irony that only the graphologists had attempted anything that looked like empirical studies, however flawed.

First, what we said in *Exorcism* when we "proclaimed" we were not going to discuss graphology was, "Note that this article does not deal with 'graphology,' the assertion that *character traits* can be divined from the analysis of handwriting."[306] And we never do discuss graphology, so defined, after the footnote in which the quoted statement appears, nor do we look at any data or cases involving graphology, so defined, in either *Exorcism* or *Science and Nonscience.* We never promised not to use books by graphologists as sources of accurate reference to independent facts, such as our citation to Huntington Hartford's (accurate) reference to Aristotle's assertion that "[j]ust as all men do not have the same speech sounds, neither do they all have the same writing,"[307] (which is the quotation Moenssens complains about),[308] or our (accurate) citation to Baldi's treatise as being the earliest modern European book on the individuality of handwriting, or our citation to Saudek to show that graphologists think they can perform identification functions; nor did we promise not to examine any data or cases or claims dealing with handwriting identification expert testimony, such as *Heller*,[309] which coincidentally happened to involve asserted experts who were graphologists.[310] We looked at asserted handwriting identification expertise globally, from whatever source derived. Yet Moenssens seems outraged that asserted identification expertise by graphologists was not excluded as an *a priori* postulate, and repeatedly accuses us of some mysterious inconsistency, or worse, for not having excluded it. Perhaps the real motive for these attacks is to be found in the last part of footnote 248, where Moenssens writes:

> The "Osbornians" among the questioned document examiners, under attack for what they perceive to be unjust and unfair criti-

[305] It is interesting to note that Robert J. Muehlberger, credited by Moenssens for his help on *Post*-Daubert *World*, gave more credit to Saudek, rightly or wrongly, in his testimony in *United States v. Martin*, than we ever did, saying that Osborn based his work on that of Saudek. *See Martin* Transcript, *supra* note 16, at 128-29.

[306] Risinger et al., *Exorcism, supra* note 2, at 733 n.13 (emphasis added).

[307] *Id.* at 734 n.14.

[308] We did establish the accuracy of Hartford's quote from Aristotle in *Science and Nonscience, supra* note 2, at 22 n.6. *See also* HUNTINGTON HARTFORD, YOU ARE WHAT YOU WRITE 43 (1973) (quoting ARISTOTLE, ON INTERPRETATION, Pt. 1).

[309] 447 N.Y.S.2d 348 (N.Y. Civ. Ct. 1981).

[310] *See* Moenssens, *Post*-Daubert *World, supra* note 1, at 307 n.248 & 303 n.231.

> cism of their profession since the 1989 *Exorcism* article, have mounted a counter-attack against Risinger and his coauthors[311] after the professors left the lofty spires of academe to descend into the pit of courtroom battle. Having become litigation antagonists, testifying as expert witnesses on the non-scientific nature of handwriting identification testimony for the defense in criminal cases against state or federal crime lab experts who take a differing view, it is only natural for Risinger and his coauthors to use any means to buttress the positions they take in litigation and direct their own ire at the "orthodox Osbornians" while giving credit to the graphoanalysts.[312]

Note that the above quotation is inconsistent with Moenssens' oft-repeated attack concerning our "apparent ignorance of the differences in training, education and professional orientation"[313] among different schools of asserted handwriting identification expertise, and his claim that we are "peculiarly unaware that a sizable number of graphologists have indeed invaded the ranks of 'handwriting experts' and have assumed the designation 'questioned document examiner'. . . ."[314] Thus, one moment he criticizes us for being unaware of the competing schools, and the next moment he criticizes us for siding with one school against the other. However, we have never "buttressed" any positions we have taken by giving any affirmative credit to graphologists, unless you count the indisputable irony noted regarding empirical research.[315] Finally, if we have directed any ire anywhere, it has been at judges and the legal system for not living up to their claimed ideology of accuracy maximization, not at any particular group of asserted experts.

27. Continuing on the "graphology" theme, Moenssens asserts in footnote 231 that we "credited the graphoanalysts' claims to competence in handwriting identification as equal to those of the Nashville participants by including their performances in the calculation of so-called 'error rates.'"[316] The reference to the Nashville participants is a little obscure, but apparently refers to the members of the Ques-

---

311 Inaccurate as it is, we believe Moenssens' article *is* that counterattack. *See supra* note 294.

312 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 307 n.248.

313 *Id.* at 303 n.231.

314 *Id.*

315 See discussion in Part I, *supra* text accompanying notes 56-58. We have never said that the attempts of graphologists at empirical research were of high quality. Their research, like Inbau's early study, was severely flawed. We have merely said that, in both instances, credit was due for at least trying, however flawed the attempt. *See* Risinger & Saks, *Science and Nonscience*, *supra* note 2, at 71 n.210.

316 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 303 n.231.

tioned Document Section of the American Academy of Forensic Sciences at the 1996 annual meeting, at which we had been invited to speak.[317] We never "included" any graphoanalysts' performances in anything. We took no steps to "exclude" such performances from the data, because there was no way to determine specifically if there were any such people in the FSF studies. However, as we have noted elsewhere, since the FSF participants were all from government labs, and since these labs are all "orthodox Osbornian" in their stated orientation, the data seem to be derived exclusively or nearly exclusively from the ranks of orthodox Osbornians.[318]

28. On page 304, Moenssens makes the following claim:

> What was certainly inexcusable in *Exorcism* was the authors' serious deficiency in legal research, in the use of inaccurate, inappropriate, out-of-date citations, and in citing authorities for propositions which they do not support. Law professors have research assistants and law review editors to prevent this from happening. Yet *Exorcism* contained innumerable errors in its citations of legal authorities, statutes, and other materials.[319]

In the same vein, in footnote 233, where Moenssens claims to document the above charges, we find the following three prefatory assertions to the three paragraphs in that footnote: "To list here all the inaccuracies in case citation and statutes is unnecessary, but to give the reader a clue,"[320] "to give just a few representative errors among many,"[321] and "[a] [c]avalier manner of using and misusing citations pervades the article."[322] Here we have the epicenter of Moenssens' McCarthyite criticism. He claims "innumerable" errors[323] (we bet we could count them), and that any examples he gives are only a small sample of all the ones he secretly knows about. How are we to respond to such claims? Well, let's look at the claimed errors he does list, remembering that he is extremely unlikely to have selected examples that are atypically trivial. Indeed, the reader would be justified in inferring that these are the most egregious examples he has, if he in fact has any more at all.[324]

---

317 The 1996 annual meeting of the AAFS took place in Nashville, Tenn., Feb. 19-24.

318 *See* note 57 *supra*.

319 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 304 (footnotes omitted).

320 *Id.* at 304 n.233.

321 *Id.*

322 *Id.*

323 *See id.* at 304.

324 Moenssens asserts that the "errors" he lists are "representative." *See id.* at 304 n.233. However, is it really believable that he would have passed up reporting a truly

a. Example 1:

> [M]ost of the statutes in Appendix 3 were inaccurate, outdated, or inapplicable. For some states, the authors cite what they say are the current statutes or rules, thus implying that none of the other states have current statutes or rules authorizing comparison of handwriting by experts, which is not true. As a matter of fact, most of the other states have current statutes or rules which should have been cited. Each instance of a failure to cite a statute or rule is an error of omission.[325]

This is truly bizarre. Appendix 3 in *Exorcism* is clearly marked "Date of Formal Acceptance of Handwriting Expert Testimony by Each State (in chronological order),"[326] and contains three columns of entry: state, date, and *instrument of acceptance* [statute, decision, court rule, etc.].[327] Since most states accepted such expertise before World War I, the cited statutes or rules (and cases) inevitably are "outdated," for goodness sake. It is true that we listed a current version of statutory or rule authority where we discovered one *in those jurisdictions where the original authority was statutory or by rule*, but as with case law, no implication that other jurisdictions did not have more recent authority, statutory or otherwise, was stated or intended. Indeed, in the text itself, we noted that by the end of the 19th century such testimony had gained general acceptance, and that by the time the article was written every state accepted expert testimony on handwriting analysis.[328] We further noted that even in those few states where we had discovered no formal authority, at the trial level "handwriting expertise was probably accepted in practice after [the Lindbergh Kidnapping case,] *State v. Hauptmann*."[329]

b. Example 2. Moenssens states that in *Exorcism* "it is stated that 'in Wyoming there is still no formal authority' admitting handwriting expert testimony . . . . In fact, Wyoming adopted the Uniform Rules of Evidence in 1978 and Rule 906(b)(s) [sic] thereof specifically authorized comparisons by experts. What's more, Wyoming case law on handwriting expert testimony goes back to 1932."[330] This is really two points. First, there is no URE 906(b)(s), or Wyoming

---

egregious error simply to maintain the typicality of his asserted sample?

325 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 304 n.233.

326 Risinger et al., *Exorcism*, *supra* note 2, at 788 app. 3.

327 *See id.* (emphasis added).

328 *See* Risinger et al., *Exorcism*, *supra* note 2, at 761-62 .

329 115 N.J.L. 412, 180 A. 809 (1935).

330 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 304 n.233 (citing Risinger et al., *Exorcism*, *supra* note 2, at 761 and State Bd. of Law Exam'rs. v. Strahan, 8 P.2d 1090 (Wyo. 1932)).

Rule of Evidence 906 anything.[331] URE 901(b)(3), like the parallel Federal Rule and Wyoming Rule, authorizes "comparison by the trier of fact or by an expert with specimens which have been authenticated" without dealing with the acceptability of any asserted expertise. As such, we did not cite such provisions unless they had been explicitly applied to handwriting, as in North Dakota and Arkansas.[332] As to *State Board of Law Examiners v. Strahan*,[333] we missed it. We would point out only that this omission was in the context of an appendix dealing with historical sources for every American jurisdiction, and, barring the reading of every case ever reported, older sources that were not then on any computer data base were easy to miss on points for which they are neither headnoted nor referenced in some other authority, such as Wigmore. Both things are true of the use of a handwriting expert in *Strahan*. For missing it, we offer the smallest of *nostra culpas*.[334]

c. Example 3. We stated that Arkansas was still "getting along nicely"[335] without handwriting experts as late as 1925, whereas Moenssens points out that "[i]n fact, Arkansas case law on handwriting expert testimony goes back at least to *Strickland v. Strickland*, 129 S.W.

---

[331] *See generally* UNIF. R. EVID.; WYO. R. EVID.

[332] Moenssens also makes the argument, rejected in *United States v. Starzecpyzel*, 880 F. Supp. 1027, 1040 n.14 (S.D.N.Y. 1995), that the language of FED. R. EVID. 901(b) requires the admission of handwriting identification expertise. *See* Moenssens, *supra* note 1, at 268 n.55. However, 901(b)(2) refers only to "nonexpert opinion on handwriting" and 901(b)(3) refers globally to "comparison by . . . expert witnesses with specimens which have been authenticated." This refers to any asserted expertise whatever, from fingerprints to earmarks, and it does not in any way prejudge the admissibility of any asserted form of expertise, novel or not.

[333] 8 P.2d 1090 (Wyo. 1932).

[334] It is also true that *Strahan* is not really authority for the admissibility of handwriting identification expertise in the sense we meant in Appendix 3, since the issue of admissibility was not dealt with in the case. Rather, the case merely notes without comment that such evidence had been admitted at trial (apparently without objection) and was part of the record. *See Strahan*, 8 P.2d at 1091. Thus, it is evidence of the *admission* of such evidence without dealing with its *admissibility*. It is generally true that any asserted form of evidence will have to have been admitted, perhaps repeatedly, at the trial level before its admissibility is authoritatively dealt with and determined in a jurisdiction. An early case fitting this pattern, for instance, is one cited by Moenssens, *Bank of Muskingum v. Carpenter's Adm'rs*, 7 Ohio 21, 31-32 (1835), where testimony by an expert is noted in the summary of evidence but not adverted to in the opinion. *United States v. Samperyac*, 27 F. Cas. 932, 940-41 (C.C. Ark. 1831) (No. 16,216a), also noted by Moenssens, does not fit this pattern, since the witnesses in that case appear to have been lay witnesses treated as sufficiently familiar with the genuine handwriting of the putative signatories to give an opinion. In any case, *Moody v. Rowell*, 34 Mass. (17 Pick.) 490 (1835), remains the earliest known American case considering the admissibility of handwriting identification expertise, as noted in *Exorcism*.

[335] *See* Risinger et al, *Exorcism*, *supra* note 2, at 761 n.136.

801 (Ark. 1910), and in 1920, the worth of handwriting expert testimony was acknowledged by the court in *Murphy v. Murphy*, 222 S.W. 721 (Ark. 1920)."[336] We did miss both cases. The *Strickland* case, like *Strahan*, was neither headnoted on the handwriting expert point, nor referenced elsewhere that we discovered, and like *Strahan*, did not really deal with the admissibility of such expertise. Half a *nostra culpa* on that. The *Murphy* case, however, was headnoted, and we concede we should have found that case and included it in the Appendix. A full *nostra culpa* on that.

d. In support of his assertion that a "[c]avalier manner of using and misusing citations pervades"[337] *Exorcism*, Moenssens says,

> For example, at the end of their discussion of the attitude of the judiciary toward document examination . . . the authors selected a particularly snide quotation from *In re Estate of Sylvestri*[338] but fail to reveal in their footnotes that the quotation is not from the decision, but from the dissenting opinion.[339]

In fact *Exorcism* says as follows:

> The current status [of handwriting identification expertise] can best be illustrated by the following quote from *In re Estate of Sylvestri*, a New York decision in sharp contrast to the thundered skepticism of *Hoag v. Wright*: "Since that rather cynical observation was made by our highest court in *Hoag*, examiners of questioned documents, as handwriting experts prefer to be called, have attained more respectable standing in the courtroom."[340]

This quote is not snide, it states the facts concerning the respectable standing of document examiners by the time of its writing in 1977, and that is the only point for which it was used in the article. However, it is true that the quotation is from a dissent (which was the only opinion filed in the matter). While not particularly relevant to the use for which the quotation was given, this fact should have been noted, and the source of the quote should not have been referred to as a "decision." So a *nostra culpa* for that one also, though it hardly shows us abusing citations to give a false picture of judicial attitudes toward handwriting identification expertise.

---

336 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 304 n.233.

337 *Id.*

338 390 N.Y.S.2d 598 (N.Y. App. Div. 1977).

339 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 304 n.233 (citing Risinger et al., *Exorcism*, *supra* note 2, at 771).

340 Risinger et al., *Exorcism*, *supra* note 2, at 771 (quoting *Sylvestri*, 390 N.Y.S.2d at 600 (Cohalan, P.J., dissenting) and citing Hoag v. Wright, 66 N.E. 579, 581 (N.Y. 1903)).

And there you have it. These are the examples which show our "serious deficiency in legal research, in the use of inaccurate, inappropriate, out-of-date citations, and in citing authorities for propositions which they do not support."[341] Most of our claimed delicts came from Appendix 3, which was not central to the Article. We missed two cases in the appendix that anyone might miss, and we missed another appendix case that we fairly ought to have found. In addition, in the main text we failed to note that one quotation was from a dissent when that was of no particular importance to the use to which we put it. We do beg the pardon of the academic community, but we have this to say: We have yet to write (or read) a formally perfect article,[342] and have little hope of ever doing so. Yet, given the number of document examiners and prosecutors who have combed this article for ten years looking for any error at all, and given that this was the worst they could find, we are reasonably well content.

---

341 Moenssens, *Post*-Daubert *World*, *supra* note 1, at 304.
342 Need we point out that this includes *Post*-Daubert *World*?