

EXORCISM OF IGNORANCE AS A PROXY FOR RATIONAL  
KNOWLEDGE:  
THE LESSONS OF HANDWRITING IDENTIFICATION  
"EXPERTISE"

D. MICHAEL RISINGER,<sup>†</sup> MARK P. DENBEAUX<sup>††</sup> & MICHAEL J.  
SAKS<sup>†††</sup>

INTRODUCTION

In 1486<sup>1</sup> the first edition of the *Malleus Maleficarum* appeared. It was a treatise on how to determine whether someone was a witch. The authors of the *Malleus*<sup>2</sup> were regarded as the leading experts of their time on this subject. They were not charlatans, nor were they weak-minded.<sup>3</sup> They were learned, analytical, and systematic, and the complex structure they set out covered 275 pages.<sup>4</sup> Their work was respected and influential in judicial proceedings for over two centuries, and inspired followers and refiners.<sup>5</sup>

---

<sup>†</sup> Professor of Law, Seton Hall University School of Law. B.A. 1966, Yale University; J.D. 1969, Harvard University.

<sup>††</sup> Professor of Law, Seton Hall University School of Law. B.A. 1965, College of Wooster; J.D. 1968, New York University.

<sup>†††</sup> Professor of Law, University of Iowa. B.A., B.S. 1969, Penn State University; M.A. 1978, Ph.D. 1974, Ohio State University; M.S.L. 1983, Yale University.

The order in which the authors are listed is the result of an agreement growing out of previous co-authorships by two of the authors, and should not be taken as an indication of relative contribution to this Article. We especially thank Carl Tobias for his trenchant editorial assistance. We would also like to thank Diane Nardone and M. Clare McLaughlin for research under pressure and James Shellow of the Milwaukee Bar for first calling the *Malleus Maleficarum* to our attention.

<sup>1</sup> The exact date is in some doubt, but 1486 appears most likely. See Summers, *Introduction to the 1928 Edition* in H. KRAMER & J. SPRENGER, *THE MALLEUS MALEFICARUM* at vii (Dover edition 1971). The title is usually rendered in English as "The Hammer of Witches".

<sup>2</sup> Heinrich Kramer and James Sprenger.

<sup>3</sup> Sprenger was a Dominican who was a distinguished scholar at the University of Cologne, holding the degree of master. He was prior and regent of his Dominican chapter, provincial of the entire German province, and was appointed general inquisitor for Germany. Kramer was a master of sacred theology who had distinguished himself in theological disputation. He was appointed jointly as inquisitor with Sprenger. See Summers, *supra* note 1, at xxxii-xxxiv.

<sup>4</sup> As printed in the Dover edition, *supra* note 1.

<sup>5</sup> See Summers, *supra* note 1, at viii ("The *Malleus* lay on the bench of every judge, on the desk of every magistrate. It was the ultimate, irrefutable, unarguable authority."). The book's standing was buttressed by the 1484 Bull of Pope Innocent VIII appointing Kramer and Sprenger inquisitors for Germany, which was published with the *Malleus* as a preface. See *id.*; see also 19 *ENCYCLOPEDIA BRITANNICA* *Witch-*

It would be an error to dismiss the social phenomenon represented by the *Malleus Malificarum* as belonging to a remote age, one we have left behind simply because we no longer believe in witchcraft. Humans in many contexts manifest a capacity to invest false expertise with undeserved reality. One can posit many reasons for this. The non-expert may wish to be spared the burdens of living in a world of unacceptably imperfect knowledge. The same goes for putative experts, except that they receive the bonuses of empowering authority and ego gratification from believing in their own capabilities—whether or not those capabilities exist. In the long run, the greatest danger of expertism is not conscious fraud or charlatanism but rather a tacit conspiracy between expert and audience to accept that which would not stand careful scrutiny because such acceptance is mutually beneficial.

This phenomenon can be seen in a wide variety of contexts, from weight control to wine tasting.<sup>6</sup> It is also not unknown in the law.

The formal recognition of the expert witness as such seems to have come late to the English common law.<sup>7</sup> Only with the rapid acceleration of science and technology beginning in the late 19th century did expert testimony become a staple of litigation.<sup>8</sup> 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

---

craft 895, 898 (15th ed. 1974).

<sup>6</sup> See the shelves of any popular bookstore for titles offering “expertise” on almost any imaginable subject. See, e.g., M. TANNY, *THE MUSCULAR GOURMET* (1988) (describing Tanny as “an expert on food and fitness . . . [who] has written the popular Muscular Gourmet column . . . and has shared life-style tips with the world’s most famous physiques and movie stars”); S. CUTHBERT, *COLORGENICS* (1987) (“Let Colorgenics show you how to harness that power to get your message across and make all your dreams come true!”).

<sup>7</sup> In the early history of English law, courts occasionally relied on a jury with specialized knowledge (for example, a jury of merchants in a commercial case) or summoned experts *sua sponte* to advise the judge. Beginning in the 16th century, the process of trial slowly evolved from an inquisitorial to an adversarial system. As it became accepted that the factfinder should consider only the testimony of witnesses who spoke from first-hand knowledge, the rule emerged excluding “opinion testimony.” The modern concept of the “expert witness” then developed as an exception to that rule. See Rosenthal, *The Development of the Use of Expert Testimony*, 2 *LAW & CONTEMP. PROBS.* 403, 406-14 (1935); see also Ladd, *Expert Testimony*, 5 *VAND. L. REV.* 414, 414-17 (1952) (discussing the rule excluding opinion testimony and its development through the inquisitorial to the adversary system in relation to the use of expert testimony). Although the exception developed in the 17th and 18th centuries, one may speculate, the “opinion rule” kept out much of what we would call expert testimony. The admissibility of “skilled witness” opinion testimony had, however, been formally recognized by 1782. See *Folkes v. Chadd*, 3 *Dougl.* 157, 159-60, 99 *Eng. Rep.* 589, 599 (K.B. 1782) (Lord Mansfield approving the testimony of a well-respected engineer as to the cause of silting in Wells Harbor, calling it “very proper evidence” and indicated that such testimony was already regularly received in practice).

<sup>8</sup> See Ladd, *supra* note 7, at 417-18.

control, and the results of judicial decisions have been spotty. The rest of this Article is about one of the spots.

Perhaps at this point the reader thinks our opening, with its invocation of the *Malleus Maleficarum*, is entertaining, but hyperbolic if applied to anything actually entered into evidence in court today. Suppose, however, a fairly common type of expert testimony existed for which the following were true:

1. No court anywhere had ever explicitly considered and passed on its claims to validity.<sup>9</sup>
2. There exist almost no studies of its claims in any academic literature.<sup>10</sup>
3. Such studies as have been conducted, published and unpublished, raise serious questions about its validity.<sup>11</sup>
4. The law has resisted requiring presentation of the asserted expertise in ways that would expose its validity problems.<sup>12</sup>

In such a circumstance, could one say that the comparison to the phenomenon of the *Malleus* was mere hyperbole? Let us now examine handwriting identification<sup>13</sup> "expertise."

<sup>9</sup> See *infra* notes 81-146 and accompanying text (Parts II and III).

<sup>10</sup> See *infra* notes 14-80 and accompanying text (Part I).

<sup>11</sup> See *id.*

<sup>12</sup> See *infra* notes 184-213 and accompanying text (Parts VI and VII).

<sup>13</sup> Note that this article does not deal with "graphology," the assertion that character traits can be divined from the analysis of handwriting. "Graphology" appears to be a somewhat newer enterprise than "chirography," the term by which handwriting identification is sometimes distinguished. Graphology was invented in mid-19th century France by a group of clerics, among them the Abbé Flandrin and the Abbé Jean-Hyppolite Michon, one of whom coined the term, and Michon's pupil Crepieux-Jamin. See H. HARTFORD, *YOU ARE WHAT YOU WRITE* 49-50 (1973).

Handwriting identification "experts" are quite clear-sighted and analytical when it comes to the weaknesses of graphology. Consider the following ironic quotes from Albert Osborn:

There are many devoted disciples of graphology throughout the world as there are of many other cults, isms and ologies.

. . . .

As practiced, at least, this alleged science has no value in identifying the author of a writing, and its exaggerations and unjustified inferences are likely to lead to loose thinking and weakening credulity. When the reports of practicing graphologists are more than mere generalities of wide application—which they usually are—they consist as a rule of positive inferences from ridiculously frail or misinterpreted data. This tendency toward exaggeration no doubt partly arises from the common tendency of devotees of various half-sciences and occult subjects to become partly self-hypnotized and self-deluded.

. . . .

The irrational human family sadly needs to be converted to a belief in the scientific method, that is, a willingness to follow truth—ignoring all

## I. DOES HANDWRITING IDENTIFICATION EXPERTISE EXIST?

When people first got the idea that the handwriting of individuals was sufficiently distinctive to enable someone to accurately assign authorship to a particular person is difficult to establish. Certainly the notion is not a new one,<sup>14</sup> and it is based on more than nothing. We have all had the experience of recognizing the author of a letter by the handwriting on the envelope alone. The best analogy for thinking about handwriting "experts" may be the practitioner of folk medicine. 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.

The last sentence ought to be startling, given the routine<sup>15</sup> admission of testimony from such asserted experts. If it fails to startle, that too is a comment on the commonly assumed relation of "expertise" to a search for truth.

The official position of the law is that expert testimony is allowed when it is rationally helpful to the trier of fact in deciding a disputed issue.<sup>16</sup> The question should be: Can the trier do as well or better in

else—wherever it may lead. We are yet a long way from this goal and man generally is still far from being a rational being.

A. OSBORN, *QUESTIONED DOCUMENTS* 442-44 (2d ed. 1929). Also note the similar sentiments of a more recent handwriting identification practitioner, James T. Miller, in Miller, *Document Examiners*, TRIAL, Aug. 1987, at 44, 44.

Testimony derived from graphology is inadmissible virtually everywhere. See, e.g., *Carroll v. State*, 276 Ark. 160, 164-66, 634 S.W.2d 99, 102 (1982) (upholding refusal to admit testimony of a "graphoanalyst" who claimed that his discipline was "an aspect of graphology" but "'much more scientific'"); *Daniels v. Cummins*, 66 Misc. 2d 575, 580-82, 321 N.Y.S.2d 1009, 1016 (Sup. Ct. 1971) (refusing to consider testimony of graphologist as to the mental competence of the person who signed a disputed deed). Graphology would seem, however, to have as much verification as some arguably admissible subjects of "expertise." See generally McCord, *Syndromes, Profiles and Other Mental Exotica: A New Approach to the Admissibility of Nontraditional Psychological Evidence in Criminal Cases*, 66 OR. L. REV. 19 (1987) (arguing that expert psychiatric testimony on such matters as witness truthfulness, compatibility of defendant's personality with crime charged, and the psychology of sexual- or spouse-abuse complainants should be admitted if a balancing of the necessity, reliability, understandability, and importance of the testimony indicates that it will assist the jury).

<sup>14</sup> Huntington Hartford quotes Aristotle as observing that "[j]ust as all men do not have the same speech-sounds, neither do they all have the same writing." H. HARTFORD, *supra* note 13, at 43 (providing no citation).

<sup>15</sup> Dean Wigmore bears a large measure of responsibility for the current state of the law. See *infra* notes 147-74 and accompanying text. The validity of handwriting identification expertise has been assumed with little question or analysis in every standard evidence treatise since Wigmore wrote. See, e.g., M. GRAHAM, *HANDBOOK OF FEDERAL EVIDENCE*, § 901.3 (1981); C. MCCORMICK, *MCCORMICK ON EVIDENCE*, § 221 at 691 (E. Cleary 3d ed. 1984); 5 J. WEINSTEIN & M. BERGER, *WEINSTEIN'S EVIDENCE* ¶ 901(b)(3)[03] (1983); 7 J. WIGMORE, *WIGMORE ON EVIDENCE* §§ 1991-2021 (Chadbourn rev. ed. 1978).

<sup>16</sup> This widely invoked principle is most importantly embodied in FED. R. EVID. 702: "If scientific, technical, or other specialized knowledge will assist the trier of fact

accuracy of decision without the benefit (or detriment) of the testimony as with it.<sup>17</sup> This is a more complicated question than it first appears, because it depends on a number of variables: how well the trier can do without the testimony, how dependable is the information provided by the testimony, and how well the trier can assess the limits of dependability of the assertions of the expert. Thus, if the trier can do pretty well without the expert but will overvalue the assertions of the expert, the expert's testimony may result in a net detriment even if it is more than randomly dependable. Central to the official rhetoric on the admissibility of any expert testimony is some demonstrated marginal efficacy for the asserted specialized knowledge above that of the average trier of fact. Such a requirement is the only thing that keeps astrology out of the courtroom. Does handwriting identification as an area of expertise<sup>18</sup> meet this requirement?

### A. *The Importance of Empirical Testing*

The question whether someone or some technique can do what it

---

to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise."

<sup>17</sup> The text reflects what we hope is a sophisticated version of the "beyond the ken of the jury" test of propriety for expert testimony. See Ladd, *supra* note 7, at 418; Slobogin, *Dangerousness and Expertise*, 133 U. PA. L. REV. 97, 135-36 (1984). Some have claimed that this test differs from the assertedly lower threshold of the "will assist the trier of fact" test reflected in FED. R. EVID. 702. See 3 J. WEINSTEIN & M. BERGER, *supra* note 15, at ¶ 702[02], §§ 702-09 to 702-10; McCord, *supra* note 13, at 71-73. We disagree. While it is easy to caricature the "beyond the ken of the jury" requirement to sound as if expert testimony would be admissible only on a point concerning which the average juror were totally ignorant, we fail to see how the proffered expert can in any rational sense "assist" the jury unless it can be shown that the expert can provide some relevant and reliable information that the average juror would not know. Interpreting "assistance" as requiring something less would, moreover, pose serious problems for judicial control. Proper control demands that the law concentrate on the issue of minimum acceptable validity, rather than assign everything to the morass of the undifferentiated balancing test ("probative value" vs. "prejudice, etc.") of FED. R. EVID. 403, as "lower threshold" proponents seem to favor. See 3 J. WEINSTEIN & M. BERGER, *supra* note 15, at ¶ 702[02], § 702-11; McCord, *supra* note 13, at 92-102.

<sup>18</sup> This note clarifies further what this Article is not about. Questioned document examiners undertake a range of examinations and utilize a variety of techniques in attempting to answer a range of questions about documents that may not be as they appear. Many of these techniques, such as ink analysis or the examination of typeface characteristics and alignments, are based on principles of applied physics and applied chemistry. Nothing in this Article deals with those techniques, or the skill of questioned document examiners in performing them. Only one broadly asserted skill and its implied biological and psychological assumptions is dealt with: the assertion that a person can learn to compare exemplars with questioned writings so that the examiner's conclusion concerning common authorship *vel non* is sufficiently accurate, and reflects a low enough rate of error, that a fact-finder can gain rationally useful information from such evidence without being misled.

purports to do is fundamental, not only for handwriting identification, but for a great variety of endeavors. Can astrology predict the course of a person's life? Can a biomedical test detect the presence of a particular disorder? Can polygraph examiners correctly classify statements as truths or lies? Can phrenology distinguish people with different personality types? How well can meteorologists predict the weather? Does a medical innovation work better, only as well as, or less well than a procedure it seeks to replace? How well can aptitude tests predict how a person will do in school or in a particular profession? Can document examiners determine whether a questioned handwriting sample was or was not made by the same person who produced a known "standard"?

The answers to these and similar questions are generally pursued by conducting empirical studies to evaluate the extent to which the claims are fulfilled. Most simply, the claim is put to the test of empirical reality. Predictions of astrologers, meteorologists, and aptitude examiners can be compared with actual outcomes. The effects of a medical treatment innovation can be compared with the effects of other treatments or no treatment (a "control"). And the classifications made by biomedical tests, phrenologists, polygraph examiners, and document examiners can be compared with some known criterion. For example, the classification (positive/negative) of a test for a certain disease can be compared with later developments in a group of patients' conditions or pathology reports.<sup>19</sup>

These kinds of empirical evaluations enable us to separate the more effective from the less effective techniques and the valid from the invalid theories. For example, Eysenck and Nias reviewed hundreds of studies of the accuracy of astrological predictions and concluded that there is no replicated and statistically significant evidence that astrologers can predict a person's personality or future from a birth chart.<sup>20</sup> In medicine, studies are done constantly evaluating and comparing various treatments; such studies have found many once-standard treatments to be ineffective or iatrogenic.<sup>21</sup> Empirical research enables practitioners to discard ineffective or harmful techniques and retain effective ones.

---

<sup>19</sup> This comparison, in essence, is the day-to-day scientific method. See K. POPPER, *THE LOGIC OF SCIENTIFIC DISCOVERY* 27, 32-34 (1968) (defining scientific propositions as ones subject to empirical testing, and therefore to potential falsification though never to absolute verification); cf. T. KUHN, *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS* (2d ed. 1970) (qualifying Popper's thesis as applicable only during periods of "normal" science when accepted axioms reign).

<sup>20</sup> See H. EYSENCK & D. NIAS, *ASTROLOGY: SCIENCE OR SUPERSTITION* (1982).

<sup>21</sup> See COSTS, RISKS, AND BENEFITS OF SURGERY 387 (J. Bunker, B. Barnes & F. Mosteller eds. 1977) (analyzing surgical procedures for which risks and costs may outweigh benefits).

There is no other way to determine which is which.

Handwriting identification might profitably be analogized to diagnostic testing in medicine or to tests used in psychology and education. Such tests are almost always subjected to validity studies.

Validity refers to the extent to which a test measures what it purports to measure.<sup>22</sup> For example, a test that purports to predict which students will do well in a pilot training program can be measured against the actual performance of pilot trainees. Some tests will inevitably measure this better than others, and the validity of the test will be determined by the degree to which it accurately separates those with the aptitude for flying (true positives) from those who lack the aptitude (true negatives). Tests with poor validity will let too many turkeys into the flight class, exclude too many top gun pilots, or both.

Most tests used in behavioral measurement have been subjected to empirical studies so that any prospective user can know, simply by looking at their "validity coefficients," the extent to which the test accurately distinguishes that which it purports to distinguish.<sup>23</sup> Similarly, in the biomedical area, virtually every test in use has a known validity—a known probability of error that can be stated with a high degree of confidence because of repeated testing of the test.<sup>24</sup>

These data permit users of these tools to know how trustworthy they are. One need not be in the position of having to rely on an inventor's or practitioner's belief concerning the efficacy of a diagnostic or predictive tool. Empirical evaluation studies permit one to place a probability statement on the conclusions a test—mechanical or

---

<sup>22</sup> As a term of art, validity refers to the actual accuracy of a test; that is, the probability that a result is accurate, and the confidence level of that probability. For an illuminating discussion of the distinction between probability and confidence and the implications of that distinction, see Cohen, *Confidence in Probability: Burdens of Persuasion in a World of Imperfect Knowledge*, 60 N.Y.U. L. REV. 385, 400-04, 409-22 (1985). Validity is sometimes contrasted with another test evaluation criterion referred to as "reliability." Both the label and the contrast are a bit confusing. First, "reliability" refers to the likelihood that two different runs of a test or process applied to the same object or event will produce the same result. It would be less confusing to refer to this likelihood as "consistency", since in everyday speech the words "reliability" and "validity" are synonymous and "consistency" more closely captures the intended meaning. Second, validity and consistency are not independent. Rather, validity is a result of a number of factors of which consistency is one. A truly inconsistent or "unreliable" test could not produce valid results, but a consistent or "reliable" test can produce invalid results because it is consistently testing some factor other than the one the tester believes is being tested. See Ennis & Litwack, *Psychiatry and the Presumption of Expertise: Flipping Coins in the Courtroom*, 62 CALIF. L. REV. 693, 697-698. (1974).

<sup>23</sup> Reliability and validity data are generally furnished by the publisher along with test materials, and the data are critically evaluated in reviews of the tests. See MENTAL MEASUREMENT YEARBOOK (J. Mitchell, Jr., 9th ed. 1985).

<sup>24</sup> See Krieg, Gambino & Galen, *Why Are Clinical Laboratory Tests Performed? When Are They Valid?*, 233 J. A.M.A. 76, 76-77 (1975).

human—offers.

Similarly, there is no reason the claimed skills and techniques of document examiners identifying handwriting cannot be put to appropriate tests in order to determine their validity. That is, how likely is it that a writing deemed genuine will in fact be genuine, and that one deemed false will be false?

### B. *The Empirical Record*

When first approaching such questions, one normally searches the relevant literature, identifies reliability and validity studies already performed, collects the study reports and reads them, and evaluates their meaning by looking at sources of potential error such as a flawed test design or errors in analysis of the results.<sup>25</sup> Usually, the literature review turns up a number of studies, and the debate then centers on how adequate they are and what inferences may be drawn from them. Not so here. Our literature search for empirical evaluation of handwriting identification turned up one primitive and flawed validity study from nearly 50 years ago,<sup>26</sup> one 1973 paper that raises the issue of consistency among examiners<sup>27</sup> but that presents only uncontrolled impressionistic and anecdotal information not qualifying as data in any rigorous sense,<sup>28</sup> and a summary of one study in a 1978 government report.<sup>29</sup> Beyond this, nothing.<sup>30</sup>

---

<sup>25</sup> An example of a result that was an artifact of the design would be a study of decisionmaker reliability that used only cases from the extremes of a distribution. This would produce inflated levels of agreement between decisionmakers. Examples of statistical errors would include failure to perform the proper (or any) significance tests or the use of the wrong correlational analysis (again creating the appearance of phenomena that are not there). See generally T. COOK & D. CAMPBELL, *QUASI-EXPERIMENTATION: DESIGN & ANALYSIS ISSUES FOR FIELD SETTINGS* (1979) (discussing the validity of several quasi-experimental designs for probing causal hypotheses); R. LIGHT & D. PILLEMER, *SUMMING UP: THE SCIENCE OF REVIEWING RESEARCH* (1984) (discussing how to review and interpret prior research on a subject); Rosenthal, *Combining Results of Independent Studies*, 85 *PSYCHOLOGICAL BULL.* 185 (1978) (summarizing methods for combining and interpreting the results of two or more independent studies testing the same hypothesis).

<sup>26</sup> See Inbau, *Lay Witness Identification of Handwriting*, 34 *ILL. L. REV.* 433 (1939). For our critique of this study, see *infra* notes 39-46 and accompanying text.

<sup>27</sup> Or, in the jargon, "reliability." See *supra* note 22 (discussing the difference between "validity" and "reliability").

<sup>28</sup> Todd, *Do Experts Frequently Disagree?*, 18 *J. FORENSIC SCI.* 455 (1973) (noting that document examiners report that their conclusions are rarely challenged by opposing expert witnesses).

<sup>29</sup> J. PETERSON, E. FABRICANT & K. FIELD, *CRIME LABORATORY PROFICIENCY TESTING RESEARCH PROGRAM: FINAL REPORT* 181-83, 194, 236-37 (1978). For our discussion of this study, see *infra* notes 48-49 and accompanying text.

<sup>30</sup> A literature search of a specialized topic risks overlooking specialized publications. To guard against this hazard, one of the authors, Professor Saks, contacted Mau-



Pause and consider what this means. From the perspective of published empirical verification, handwriting identification expertise is almost nonexistent. Nor is it an area, like practical plumbing, in which the results of correct or incorrect practice would be obvious to bystander and practitioner alike.<sup>31</sup> Rather, it is exactly the kind of situa-

---

reen Casey Owens, Chief Document Examiner at the Crime Laboratory of the Chicago Police Department and then president of one of the associations for such practitioners, the American Society of Questioned Document Examiners. She replied with a bibliography published by her society and a cover letter identifying the primary books and journals on which, she stated, document examiners rely. As a result of her aid, supplemented by further research and cross references, the authors reviewed the literature reflected in Appendix 1 of this article. The vast majority of handwriting "studies" reported in these publications consist of anecdotal observations, hypothesis creation, and speculation.

<sup>31</sup> There is another avenue of development that handwriting identification expertise might have taken, but it has never been explored beyond a few tentative steps. One might develop a system that, first, contains a reliable criterion for dividing parts of an example of handwriting into appropriate standardized individual units; second, contains a reliable method of classifying those units; and third, develops data about the statistical incidence of those units in various populations of writers, their correlates, and the effect of attempts at disguise upon them. Were that done, the process would in theory be freed of much of the subjectivity that is its current plague. Nothing of the sort has been accomplished, however, even though leaders in the questioned document community such as Ordway Hilton have recognized the desirability of such an approach. See Hilton, *How Individual Are Personal Writing Habits?*, 28 J. FORENSIC SCI. 683, 684-85 (1983). There exist demanding problems in developing a unitary classification system for so curvy a thing as writing, and then selling it to the whole world. A simple classification system adequate for indexing but inadequate for statistical purposes was proposed by Abbey and Lee in 1923. See Abbey & Lee, *Abbey-Lee Handwriting Classification*, 13 J. CRIM. L. & CRIMINOLOGY 140 (1922). This system gained widespread but by no means universal adoption for indexing police exemplar files. In 1954, a similar system was proposed without reference to Abbey & Lee by Theodora LeH. Smith. See Smith, *Six Basic Factors in Handwriting Classification*, 44 J. CRIM. L. & CRIMINOLOGY 810 (1954). In 1959, Orville Livingston published a criticism of the Abbey-Lee system along with a proposal for a new classification system. His system, unfortunately, also was inadequate for statistical purposes. See Livingston, *A Handwriting and Penprinting Classification System for Law Violators*, 49 J. CRIM. L. & CRIMINOLOGY 487, 487 (1959). Although it is very difficult to create a body of statistical research without a standardized scheme of classification, a few "proto-statistical" studies have been performed. See, e.g., Harris, *Disguised Handwriting*, 43 J. CRIM. L. & CRIMINOLOGY 685 (1953) (attempting to determine which changes of writing were most common in attempted disguise by examining disguises produced by a student exercise); see also Harris, *How Much Do People Write Alike?: A Study of Signatures*, 48 J. CRIM. L. & CRIMINOLOGY 647, 647 (1958) (comparing and contrasting signatures on voter registration lists produced by different people with the same last name). Harris presents photographs of signatures showing so little variation that determining whether they are by one or by different hands is a daunting task. Even more interesting are the signatures Harris reports he found but did not photograph: "So many of these signed [sic] lacked individuality and looked alike that they were not worth photographing." *Id.* More recently, there have been several more formal attempts at statistical study, see Totty, Hardcastle & Dempsey, *The Dependence of Slope of Handwriting upon the Sex and Handedness of the Writer*, 23 J. FORENSIC SCI. SOC'Y 237, 237 (1983), and a much more ambitious project, Eldridge, Nimmo-Smith, Wing & Totty, *The Variability of Selected Features in Cursive Handwriting: Categorical Measures*,

tion in which it is easiest to fool oneself and others. If handwriting expertise were offered for the first time today with this published record as its foundation, courts would almost certainly reject it.

As it turns out, the published record does not reflect all of the empirical studies that have been done. 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. Handwriting identification was included in the original studies in 1975; that test is the one reported in the 1978 government publication referred to above.<sup>32</sup> A permanent yearly testing program was begun in 1978,<sup>33</sup> and a handwriting component was added to the program in 1984. Thus there are now four unpublished FSF studies for which we have obtained the data: 1984,<sup>34</sup> 1985,<sup>35</sup> 1986,<sup>36</sup> and 1987.<sup>37</sup> These FSF studies, the 1978 study, and the 1939 study by Inbau,<sup>38</sup> constitute the entire corpus of empirical data on handwriting identification, so far as we have been able to discover. We now turn to an examination and evaluation of that corpus.

### 1. The 1939 Inbau Study

The earliest study was done in 1939 by Fred Inbau, a colleague of

---

24 J. FORENSIC SCI. SOC'Y 179, 179 (1984). However, even if a statistical approach might someday provide dependable and useful information in legal proceedings, experience with other such enterprises suggests it will be decades before an adequate system and a sufficient data base are in place, especially given the current lack of academic base for the enterprise. *See infra* note 183. It is no coincidence that the beginnings noted here also represent the first interest in the field shown by academic researchers. *See generally* Davis, *Forensic Handwriting Research at Birmingham University*, 23 J. FORENSIC SCI. SOC'Y 251, 251 (1983) (describing work being done in the English Department of Birmingham University on forensic handwriting analysis).

<sup>32</sup> *See* J. PETERSON, E. FABRICANT & K. FIELD, *supra* note 29.

<sup>33</sup> Lucas, Leete & Field, *An American Proficiency Testing Program*, 27 FORENSIC SCI. INT'L 71, 74 (1985).

<sup>34</sup> *See* COLLABORATIVE TESTING SERVICES, INC., CRIME LABORATORY TESTING PROGRAM, REP. NO. 84-7, QUESTIONED DOCUMENT ANALYSIS (1984) (on file with *The University of Pennsylvania Law Review*) [hereinafter 1984 FSF STUDY]

<sup>35</sup> *See* COLLABORATIVE TESTING SERVICES, INC., CRIME LABORATORY TESTING PROGRAM, REP. NO. 85-8, QUESTIONED DOCUMENT ANALYSIS (1985) (on file with *The University of Pennsylvania Law Review*) [hereinafter 1985 FSF STUDY]

<sup>36</sup> *See* COLLABORATIVE TESTING SERVICES, INC., CRIME LABORATORY TESTING PROGRAM, REP. NO. 86-5, QUESTIONED DOCUMENTS ANALYSIS (1986) (on file with *The University of Pennsylvania Law Review*) [hereinafter 1986 FSF STUDY]

<sup>37</sup> *See* COLLABORATIVE TESTING SERVICES, INC., CRIME LABORATORY TESTING PROGRAM, REP. NO. 87-5, QUESTIONED DOCUMENTS ANALYSIS (1987) (on file with *The University of Pennsylvania Law Review*) [hereinafter 1987 FSF STUDY]

<sup>38</sup> *See* Inbau, *supra* note 26.

Wigmore's at Northwestern University Law School.<sup>39</sup> This study was not planned as a study of document examiner validity but was intended and designed as a study of the ability of non-experts to identify handwriting.<sup>40</sup> The addition of document examiners was an afterthought, which may help to explain a number of the study's weaknesses.

Inbau had eleven of his colleagues on the faculty at Northwestern write their own signatures on five separate cards. For each writer he then had prepared a "spurious" signature, one made by someone other than the actual author in the "forger's" natural hand; a freehand forgery by a person with no experience or known skill at forgery (made by someone looking at an actual signature while trying to copy it); and a traced forgery (made by someone tracing an actual signature).<sup>41</sup> Inbau then tested the ability of several samples of people to distinguish an actual signature from the various inauthentic signatures. He presented one genuine signature, one "spurious" signature, one freehand forgery, and one tracing to eight everyday people, seven bank employees, and three document examiners.<sup>42</sup> While the results for the document examiners were marginally better than the other groups numerically, none of the differences was statistically significant.<sup>43</sup>

Though this test failed to produce any evidence of meaningful difference between document examiners and others, methodological defects in the study prevent it from being used as a basis to draw virtually any conclusion. The obvious first problem is the small sample sizes, especially the small number of document examiners. Inbau reported that he had great difficulty getting professional document examiners to participate in the study.<sup>44</sup> In the end, he had to settle for only two practicing examiners and a third person whom Inbau considered to be capable even though not a practicing document examiner. Thus, Inbau's sample of document examiners was both small and skewed by a self selection that probably eliminated those who were insecure about putting their

---

<sup>39</sup> *Id.* at 433, n.1.

<sup>40</sup> *See id.* at 436.

<sup>41</sup> *See id.* at 434.

<sup>42</sup> *See id.* at 439.

<sup>43</sup> Inbau provides a raw data table. *See id.* at 439. Inbau notes that the data he collected revealed no statistically significant differences between the layman's and bank tellers' ability to identify genuine signatures. *See id.* at 436 n.7, 440 n.10. Our own Kruskal-Wallis analysis of variance revealed that the differences among the three groups were not statistically significant. ( $H=5.06$ ,  $df=2$ ,  $p < .05$ ). *See S. SIEGEL, NONPARAMETRIC STATISTICS FOR THE BEHAVIORAL SCIENCES* 184-94 (1956) (explaining the methodology and meaning of Kruskal-Wallis analyses).

<sup>44</sup> 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." Inbau, *supra* note 26, at 440 n.11.

abilities to a test, but who nevertheless were willing to testify in court as experts.

The second problem is more protean: The task selected for the test may not have been of meaningful difficulty. 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 known to produce, the fifth is a professional calligrapher, and the note is in flowing calligraphy, then identifying the writing would be easy for anyone. On the other hand, some identification tasks may be beyond anyone's skill. If one were to select a person whose handwriting was especially nondescript and easy to copy, and set five skilled and dexterous forgers to work for a month mastering that person's writing, then the results might be indistinguishable by any living person. In between these two extremes lies a broad variety of circumstances and tasks. Tests must be designed carefully to present discriminations of meaningful difficulty and variety. Only results from such tests could begin to paint a picture of what both lay people and experts can and cannot do and when, if ever, "experts" can perform significantly better than the average person. If the task is too hard (like a brain transplant), expert and non-expert alike will fail and one might conclude that no expertise exists; if the task is too easy (like cleaning and dressing a scraped knee), both may complete the task successfully and again one might conclude that no expertise exists. Only tasks in the middle range can reveal whether expertise exists and if so what its bounds are.

It is not as easy to design such tests as one might assume at first blush. Take the apparently easiest task in the Inbau study: identifying the genuine signature and the "spurious" signature in someone else's everyday hand. Looking at the test cards used in the Inbau study,<sup>45</sup> it is clear that few of the genuine signatures at all resemble any of the spurious signatures. If the question is put "assuming both signatures were written normally with no attempt at disguise, is signature #2 in the same handwriting as signature #1," everyone, lay and expert, should be 100% correct. If we ask instead whether signature #2 was written by the same *person* as signature #1, however, the answers may represent a different meaning even if they do not convey different re-

---

<sup>45</sup> See Hilton, *The Detection of Forgery*, 30 J. CRIM. L. & CRIMINOLOGY 568, 576, 579, 582, 586-87, 590, 594 (1939) (containing reproductions of Inbau's test cards).

sults. The lay people might take this question to be identical to the first one, but the experts might examine for the possibility that one of the signatures is a second or disguised hand of the person who wrote the other or that both are disguises from a third "normal" hand. Both groups may still get the same percentage correct, but hidden in that equivalence is the fact that they performed different tasks. If one asks "Is signature #2 a disguised signature of the person who wrote #1?," this question may have a substantial effect on the way the lay people approach the task (and some of the experts, for that matter) and change the results dramatically. In the case of the Inbau study, the nature of the instructions given to the participants, and the test conditions, are almost entirely unknown. We don't even know if the tests were conducted in an acceptably blind manner.<sup>46</sup>

The point here is not to be overly critical of Inbau. His attempt was at least a start; if anything further had been done in the following three decades we might now know a good deal more about the validity of this asserted expertise. The point is rather that good tests that give meaningful results applicable to the situations that arise in legal proceedings are not easy to design, and there would have to be a whole variety of them developed and repeatedly administered before any real verification would begin to emerge.

## 2. The Forensic Science Foundation Studies

Now let us turn to the unpublished FSF studies. Some preliminary observations are in order before describing the studies themselves. First, these tests were developed as proficiency tests for specialists by a group that assumed there was such a thing as competency. That is, while they were willing to believe there might be some incompetent practitioners, they assumed that there were competent practitioners whose results would be dependably valid, and who would always outperform you or me. Hence, the tests were not presented to control

---

<sup>46</sup> What we might infer of the conditions of test administration from some of Inbau's description is hardly encouraging:

Two additional professional examiners agreed to assist, but, unfortunately, one of these had been inadvertently informed (by a colleague of the writer's) prior to the test that each of the signature sets contained only one genuine specimen—and for this reason the expert's results, which were otherwise perfect, had to be discarded. The other examiner, after retaining the exhibits over a period of six months without submitting a report, returned them with his apologies. . . . He submitted a semblance of a report, but it was lacking in specific statements which could [permit a test of his correctness or incorrectness].

Inbau, *supra* note 26, at 440 n.11.

groups of non-"experts" to determine if the problems presented were too easy. Second, the people to whom the tests were given knew they were being tested, so the problem of special effort devoted to competency tests was present.<sup>47</sup>

a. *The 1975 test*<sup>48</sup>

In the 1975 test, researchers sent a letter composed of both typewriting and handwriting to the participating examiners along with four exemplars written by four different people. Examiners were asked to determine whether any of the "suspects" executed the handwriting on the questioned document. In actuality, one of the four "suspects" had written the letter.<sup>49</sup> Of the seventy-four responding laboratories:

- 66 (89%) correctly identified the suspect who wrote the questioned letter.
- 1 (1%) reported partially correct and partially incorrect results<sup>50</sup>
- 4 (5%) said they could not make any conclusion from what had been submitted.
- 3 (4%) identified the wrong person.

---

<sup>47</sup> The alternate possibility exists that, these being merely tests, the document examiners gave them less attention than they give to the rest of their work. The amounts of time devoted to the proficiency tests by the examiners does not, however, appear to be insignificant: the mean time they spent on the task was, in 1984, 2.99 hours; in 1985, 3.25; in 1986, 2.84; and in 1987, 3.72. *See infra* Appendix 2. It must be noted here that FSF's Proficiency Advisory Committee disavows these tests as representative of the level of performance of 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?

<sup>48</sup> J. PETERSON, E. FABRICANT & K. FIELD, *supra* note 29.

<sup>49</sup> The question asked in the 1975 test was, in pertinent part: "You are asked to determine which (if any) of the suspects prepared the handwriting on the threatening letter" *Id.* These instructions use the word "suspect." In real cases, the very act of submitting samples to a crime lab is an assertion that there is other evidence incriminating the suspects, unless the process of submitting samples "line-up fashion" with other masking exemplars is or becomes standard. This problem is of course magnified if there is only one "suspect." Even then, to be submitted in proper lineup fashion, the known false exemplars ought not to be too easily discarded, and should be in two independent groups to prevent the "forcing" response of picking the most likely example in the group as the confirmation of the implied assertion, resulting from submission, of the existence of other evidence.

<sup>50</sup> That is, they identified the correct suspect as having written part of the note, but incorrectly attributed another part of the note to an incorrect suspect. *See id.* at 236.

b. *The 1984 test*<sup>51</sup>

In the 1984 test, three handwritten letters containing bomb threats were said to have been received by the news media and followed by bombings by a terrorist organization. Forty-one participating laboratories employing document examiners were sent the three letters plus twelve pages of known handwriting samples (two pages for each of six suspects).<sup>52</sup> Examiners were asked to determine whether all of the questioned letters were written by the same person, and whether any of the questioned letters were written by any of the authors of the known writings.<sup>53</sup> Two of the letters were written by one person, whose writing was not submitted, and the third letter was written by one of the suspects whose exemplar was in his normal hand but who in writing the questioned letter attempted to simulate the writing in the other two letters.<sup>54</sup> Only twenty-three of the forty-one laboratories submitted reports:

- 17 (74%) caught the different authorship of the third letter.
- 6 (26%) said erroneously that all the letters were written by the same person
- 23 (100%) failed to recognize the author of one of the questioned letters among the known exemplars. Thus they were all correct concerning letters 1 and 2 and all wrong concerning letter 3. Every examiner failed to recognize the author of one of the questioned letters among the exemplars.<sup>55</sup>

c. *The 1985 test*<sup>56</sup>

Twelve checks all bearing a signature in the same name were sent to forty-two participating laboratories. They were asked to determine

---

<sup>51</sup> See 1984 FSF STUDY, *supra* note 34.

<sup>52</sup> See *id.* at 8 (Appendix).

<sup>53</sup> This is actually the end result of a number of sub-questions for which we have no data: Was the writer of letter 1 the writer of letter 2? Was the writer of letter 1 the writer of letter 3? Was the writer of letter 2 the writer of letter 3? Did any of the six write letter 1 (six questions)? Did any of the six write 2 (six questions)? Did any of the six write 3 (six questions)? Thus, to perform the test, answers to twenty-one separate questions are required. (Although they are linked; for example, a decision that two or three of the letters were written by the same author would consolidate several of these questions). Some may be hard and some not.

<sup>54</sup> See *id.* at 3.

<sup>55</sup> See *id.* at 2.

<sup>56</sup> See 1985 FSF STUDY, *supra* note 35.

which if any of the signatures were made by the same person.<sup>57</sup> In actuality, two of the twelve had been signed by the same person. A third was a freehand forgery by a person of no known skill or experience as a forger. A fourth was a tracing. The others were signed by different people in their own hands.<sup>58</sup> Note the similarity of this test's design to that of the 1939 Inbau test.<sup>59</sup> Only thirty-two of the forty-two laboratories returned the 1985 test. Of those,

- 13 (41%) gave correct results.
- 2 (6%) of responses were incorrect in attributing one of the forgeries to the real repeat signatory.
- 10 (31%) said they were unable to reach conclusions.
- 7 (22%) were substantially wrong.<sup>60</sup>

d. *The 1986 test*<sup>61</sup>

The scenario for this study was that police stopped a car with three known occupants. In the car they found incriminating evidence, including a hand-printed note used in the commission of the crime. Samples of printing were obtained from the three occupants of the car.<sup>62</sup> The test materials included two samples taken from an actual case and written by the same author, one submitted as the "crime note" and one as an exemplar. Two other exemplars were added for the test, one in a document examiner's normal printing and one produced by a document examiner attempting to simulate the printing of the crime note.<sup>63</sup> The subjects' task was to determine which, if any, of the writers of the exemplars wrote the crime note. The materials were requested by forty-eight laboratories. Thirty-one responded:

- 4 (13%) gave correct answers.
- 3 (9%) were partially correct in that they said none of the authors of the exemplars wrote the crime note, when in fact two had not, but one had.
- 10 (32%) were unable to reach conclusions.
- 14 (45%) gave incorrect answers, mostly by assigning authorship to the forger.<sup>64</sup>

---

<sup>57</sup> See *id.* at 1.

<sup>58</sup> See *id.* at 2.

<sup>59</sup> See *supra* notes 39-44 and accompanying text.

<sup>60</sup> See 1985 FSF STUDY, *supra* note 35, at 2-3.

<sup>61</sup> See 1986 FSF STUDY, *supra* note 36.

<sup>62</sup> See *id.* at 16 (Appendix).

<sup>63</sup> See *id.* at 5.

<sup>64</sup> See *id.* at 6-10.



e. *The 1987 test*<sup>65</sup>

Because of complaints from document examiners that prior tests were too difficult, the Proficiency Advisory Committee decided to make the 1987 test easy. According to the report, "[t]his test was designed to be a relatively easy and straightforward test, because of complaints about previous test design. All the writings in this test were natural and free of disguise."<sup>66</sup> This test involved a single questioned extortion note written by one suspect. Both request and non-request known exemplars from four suspects were provided to the participants. After comparing the questioned and the known, examiners were to offer an opinion concerning which, if any, of the known exemplars was written by the same person who wrote the questioned note.<sup>67</sup> Of fifty-five laboratories requesting materials, thirty-three returned reports:

- 17 (52%) gave correct answers.
- 1 (3%) incorrectly eliminated the correct suspect.
- 0 (0%) incorrectly identified an innocent suspect.
- 15 (45%) were unable to reach a conclusion.<sup>68</sup>

What do all five FSF studies taken together suggest? A rather generous reading of the data would be that in 45%<sup>69</sup> of the reports forensic document examiners reached the correct finding, in 36%<sup>70</sup> they erred partially or completely, and in 19%<sup>71</sup> they were unable to draw a conclusion. If we assume that inconclusive examinations<sup>72</sup> do not wind up

---

<sup>65</sup> See 1987 FSF STUDY, *supra* note 37.

<sup>66</sup> *Id.* at 1.

<sup>67</sup> See *id.* at Appendix.

<sup>68</sup> *Id.* at 4-5. The FSF Proficiency Advisory Committee Comments on these results are telling: "Considering the relative simplicity of this test, the overall results were disappointing because of the large number of inconclusives and the one false elimination. The fact that 17 of 33 labs did report all correct answers is some consolation." *Id.* at 2.

<sup>69</sup> This percentage is the mean of the percentages of correct answers in each of the tests (89% in 1975, 74% and 0% in 1984, 41% in 1985, 13% in 1986, and 52% in 1987). An even more expansive reading of the data would calculate instead the percentage correct of all the answers given in all the years taken together. Of the 193 total responses, 117 or 61% were substantially correct (66 in 1975, 17 in 1984, 13 in 1985, 4 in 1986, and 17 in 1987). This reading is more than generous, however; it is extravagant. It, in effect, weights the results of each year's test by the number of responses each test drew. Since that number appears to be inversely related to each test's difficulty, such a weighting would simply exacerbate the problems of sample bias discussed *infra* note 76 and accompanying text. Moreover, it would unbalance the FSF series' scheme of testing a variety of handwriting analysis skills.

<sup>70</sup> The mean percentage of erroneous conclusions. (1975: 5%; 1984: 26% and 100%; 1985: 28%; 1986: 55%; and 1987: 3%).

<sup>71</sup> The mean percentage of inconclusive reports (1975: 5%; 1984: 0% and 0%; 1985: 31%; 1986: 32%; and 1987: 45%).

<sup>72</sup> There is at least some anecdotal historical evidence that prosecutors can negoti-

as testimony in court, and omit the inconclusive reports, and remain as generous as possible within the bounds of reason, then the most we can conclude is this: Document examiners were correct 57% of the time and incorrect 43% of the time.<sup>73</sup>

But let us turn to more meaningful readings of the aggregate data. The pilot test in 1975 may have been unrealistically easy, like a line-up with four beefy white policemen and a skinny black person. Did this task present any real difficulty at all? There is no way of knowing whether a group of lay persons would have done any less well, since none was tested. Omitting the 1975 data, the examiners were correct 36% of the time, incorrect 42%, and unable to reach a conclusions 22% of the time.<sup>74</sup> Even these results are biased in favor of accuracy because of the intentional ease of the 1987 test.<sup>75</sup> Disguised handwriting fooled them all and forged printing fooled two-thirds of those who hazarded an opinion about it.

Now consider the effect on the aggregate results of the laboratories that requested test materials but did not return them.<sup>76</sup> More likely than not, these non-respondents bias the results further in favor of correct conclusions. Some of the non-responding labs, no doubt, did not even perform the tests due to the press of daily business. But some others very likely performed the tests and then did not return their reports. Assuming that an examiner who has worked on an answer and then decides not to return it has serious doubts about its accuracy, then

---

ate with forensic scientists to turn inconclusive or even negative reports into something that sounds inculpatory in court. This is what apparently happened with the ballistics in the government's case against Sacco and Vanzetti. See G. JOUGHIN & E. MORGAN, *THE LEGACY OF SACCO AND VANZETTI* 15-16 (1976).

<sup>73</sup> This capriciousness contrasts with claims by document examiners that their field's performance is just shy of perfect. See Todd, *supra* note 28, at 458-59 (stating that "it appears safe to answer a categorical 'No' to the question whether document examiners frequently disagree").

<sup>74</sup> Of the 84 responses received with asserted conclusions in the 1984 through 1987 tests, only 34 (40%) were completely accurate (0 in 1984, 13 in 1985, 4 in 1986, and 17 in 1987).

<sup>75</sup> Recall that the 1987 test was designed to be "easy and straightforward." See *supra* note 66 and accompanying text. Yet even then the laboratory directors who supervised the program were disappointed that only about half of the examiners reporting could produce a correct answer. See *supra* note 68.

<sup>76</sup> For the 1975 study, we do not know how many laboratories were sent handwriting materials, but we do know that approximately 250 of them across North America participated in the larger study (the wider range of forensic science proficiency testing, including handwriting). See J. PETERSON, E. FABRICANT & K. FIELD, *supra* note 29. We also know that 74 returned reports for the handwriting test. The Forensic Science Foundation's regularized series of proficiency tests sent materials only to labs that requested them. In 1984, 18 of the labs that ordered test materials did not return reports of their examinations; in 1985, 10 did not return reports; in 1986, 17, and in 1987, 22.

the sample of respondents is composed of an unrepresentatively large proportion of those who obtained—or at least think they obtained—correct answers.

If a correct answer consists of a report containing correct conclusions returned pursuant to requested and submitted test materials, then of the total submissions to laboratories in the 1984 through 1987 tests, only 18% gave wholly accurate responses (without the 1987 test the figure drops to 13%).

Finally, consider the possible effect on any aggregate conclusions of the fact that, of the more than 250 police laboratories that perform handwriting examination (not to mention a large number of private practitioner document examiners), only a fraction even ordered test materials in the first place. It is at least arguable that, by self-selection, the sample is inherently biased in favor of the more conscientious and capable practitioners to begin with. If this is true, the reported results would overstate the accuracy of the handwriting examination field generally.

The 1984, 1985, and 1986 tests presented examiners with a variety of challenges. The results should provide anyone with cause for concern. The examiners who returned reports on the analysis disagreed among themselves a good deal of the time, suggesting limited reliability, and many of the opinions offered were incorrect, suggesting limited validity.

In addition, the studies failed to reveal that certification or experience enhanced accuracy. The 1987 Proficiency Advisory Committee Comments state that "[a]s usual, there were no correlations between right/wrong answers and certification, experience, amount of time devoted to document examination and length of time spent on this test."<sup>77</sup> Consider what this independence means for a court's likely assumptions about whether to admit a proffered expert and for the weight a factfinder is expected to give such testimony. A court is likely to assume that an examiner who is certified, who has been on the job for many years, whose caseload is nothing but document examination, and who has spent a lot of time examining the evidence, is especially likely to have something useful to say to a jury. Yet these data provide no support for these assumptions. Examiners who are uncertified, have little experience, work on document examination only part time, and spend little time on the particular document, are just as likely to be right as someone with more impressive qualifications. Does any of this suggest

---

<sup>77</sup> 1987 FSF STUDY, *supra* note 37, at 2.

the existence of expertise?<sup>78</sup>

These are the sorts of findings about the nature and limits of asserted handwriting identification expertise about which both document examiners and the courts need to know but which could not have been known before such studies were undertaken. Though they are not without flaws, these studies represent a step toward systematic and scientific evaluation of the claimed capabilities of this asserted expertise. Perhaps some of the considerably larger number of needed tests yet to be designed and administered would show document examiners faring better, but on the present record we must say that the underpinnings of the "expertise" have degenerated from no data to negative data.

Finally, we cannot emphasize too strongly that from the viewpoint of the law each of these studies suffers from a major omission: the absence of a control or comparison group of lay test-takers.<sup>79</sup> If a jury can compare handwriting no worse than proffered "experts," then the expertise does not exist. For any given task, the level of performance of professional document examiners may be no better than that of laypersons. Indeed, lay persons might perform some tasks consistently *better* than "experts." While such superiority may seem intuitively improbable, it remains a logical possibility and one not without analogues in other areas.<sup>80</sup> For now, the kindest statement we can make is that no

---

<sup>78</sup> Although the reports state that no relationships exist between accuracy and several of the variables that one might expect to be related to accuracy, they present no analyses to support that conclusion. Fortunately, the reports provided the raw data, permitting us to carry out these statistical analyses ourselves. *See infra* Appendix 2.

<sup>79</sup> Bear in mind that these studies were never intended to answer the question that courts need answered; they simply come closer to answering it than anything else that exists. Studies with a design capable of directly answering the central question—the performance of handwriting identification experts versus non-experts—have never been undertaken.

<sup>80</sup> One example occurs in psychiatric predictions of dangerousness (although in fairness we must point out that while some psychiatrists and psychologists mount the witness stand to offer their predictions, their respective professional associations disavow the existence of such special expertise.) *See* Dershowitz, *Psychiatry in the Legal Process: A Knife That Cuts Both Ways*, TRIAL, Feb/Mar 1968, at 29, 32 (citing studies reporting that prison staff were better able to predict dangerous behavior than psychiatrists—perhaps because psychiatrists, whose predictions of dangerousness result in commitment, rarely get the feedback to know whether their predictions are correct); *see also* D. HOGAN, THE REGULATION OF PSYCHOTHERAPISTS 114 (1979) (reporting studies "indicating that theoretical and technical knowledge is largely irrelevant to effectiveness.")

Finally, it should be noted that experts may also become *less* dependable over time. Progress is not inevitable, particularly if conditions change. For instance, one of the handwriting identification factors pointed to by Osborn and others as most distinctive was "shading," variations in line width that resulted from the use of nib pens. *See* A. OSBORN, *supra* note 13, at 126-40. With the disappearance of nib pens, shading is no longer significant. *See* Livingston, *supra* note 31, at 488 (referring to the "practically obsolete shading factor"); *see also* S. FAY, L. CHESTER & M. LINKLATER, HOAX

available evidence demonstrates the existence of handwriting identification expertise.

If you have read this far, two questions may well spring to mind: How did the law come to accept such expertise, and what should be the response of courts when they are alerted to the absence of verification for handwriting identification expertise? The latter question is the quickest answered: If courts mean what they say about the nature and function of expert testimony and when it should be admitted, they will cease to admit it. That will almost certainly not happen, however, and we will hold off a consideration of the implications of *that* until we address the first question: How did the law come to accept the existence of handwriting identification expertise?

## II. THE HISTORY OF HANDWRITING EXPERTISE IN COURT

Documentary evidence is important in a great variety of cases. While this statement may have been slightly less true a few hundred years ago, documents have always been potentially significant in much litigation.

The relevancy of documents<sup>81</sup> is often conditioned on their authorship or origin, or on their acceptance or ratification. Lawyers have long encountered difficulties in generating evidence to counter denials of authorship or of acceptance of documents.<sup>82</sup> Various ways of reducing this problem were early worked out for documents such as written contracts, bonds, wills, or deeds, when controversy could be anticipated and when authorship or ratification could readily be tied down at the moment of writing.<sup>83</sup> Witnesses to the drafting or signature, and the application of complex seals, were chief among the strategies adopted. Indeed, the problem was looked upon as so great that absent witnesses,

---

131 (1972) (describing the variations that occur with the use of different writing instruments and the quality of forgery that can be expected). If shading were critical to valid identification, the "expertise" could move from valid to invalid with the disappearance of shading.

<sup>81</sup> When the documents are typewritten, typeprinted, or otherwise mechanically produced, or when the acceptance or ratification is not by signature, the issues must be resolved by reference to evidence not directly dealt with in this article. When all or part of a document is handwritten or when it bears a handwritten signature, the kind of evidence dealt with in this article becomes potentially important.

<sup>82</sup> "There is nothing so difficult to put beyond question as the fact that a particular instrument, which the witness has not seen to be signed, was signed by a particular person." Kellogg, *Proof of Handwriting*, 12 CENTRAL L.J. 507, 507 (1881) (quoting Lord Coleridge).

<sup>83</sup> See 3 J. BENTHAM, RATIONALE OF JUDICIAL EVIDENCE 587 (1827) (labelling this phenomenon "preappointed evidence" and affirming its purpose as "prevention of spurious or falsified instruments").

some documents were without legal effect.<sup>84</sup> There always existed, however, a variety of situations in which such mechanisms were unlikely expedients. Anonymous or unsigned documents were potentially important in many criminal cases, and the same could be true in regard to civil libel or conspiracy. Furthermore, the promises of *assumpsit* were regularly reflected in unsigned or unsealed documents.<sup>85</sup> So the authentication of handwritten documents as to authorship or origin—the provision of sufficient proof to justify presenting them to the trier of fact—early became an issue.

A modern analysis of the fact patterns presented by many important early cases on this issue would probably find that the reasonableness of a jury's finding of authorship or ratification was supported primarily by circumstantial evidence. Thus, in *Algernon Sydney's Case*,<sup>86</sup> the prosecution relied heavily upon an incriminating but unsigned letter. The letter had hardly blown in from a busy London street. It had been found on Sydney's desk, in his home, among a number of other papers.<sup>87</sup> While such circumstance may not prove the required authorship beyond doubt, it provides a rather significant starting point in assessing the probabilities. However, the analytic tendency seems to have been to view the issue of authorship formally as in some sense independent and divorced from external circumstances. Courts throughout the 17th and 18th centuries generally followed a similar approach. They lamented the absence of a witness who had watched the writing being written. They then considered what might be acceptable in the absence of such a witness.<sup>88</sup>

The structure of this approach created a problem that has continued to plague handwriting authentication down the centuries. If too much supplemental evidence is required to authenticate a document formally, independent of the particular circumstantial evidence available in the case, formal denial of authorship or origin may result in a formal failure of proof even when circumstantial evidence indicates an overwhelming likelihood of authorship. Conversely, if too little is formally required, unjustified decisions on authentication may result even

---

<sup>84</sup> The most common example is wills, which many jurisdictions even today require to be witnessed. See 4 J. WIGMORE, *supra* note 15, § 1290, at 704.

<sup>85</sup> Lack of formal requirements, even unto the enforceability of an oral promise, was of course the hallmark of English contract law as it developed within the action of *assumpsit*.

<sup>86</sup> 9 How. St. Tr. 818 (1683). The writing complained of allegedly constituted evidence of treason by its author.

<sup>87</sup> See *id.* at 854 (testimony of Sir Philip Lloyd)

<sup>88</sup> See *supra* note 82 (quoting a statement of Lord Coleridge) for a reflection of this general approach.

absent strong supplemental circumstances.

Because they failed to focus on circumstantial evidence, none of the historical authentication standards dealt directly with the main problems inherent in the authorship question. However, it should also be noted that there was one joker absent from the game until a fairly recent time: expert testimony. As Wigmore observed, as late as the early 19th century, "the idea of expertism in handwriting was . . . a novel one."<sup>89</sup> In early cases, materials presented to the factfinder arrived without the extra bang, rational or irrational, that results when the legal system allows testimony to be presented wrapped in the mantle of expertise.

Instead, the contours of the "pre-expertise" debate focused on the following issues, though in no particularly systematic way: Was a factfinder allowed to compare by physical juxtaposition the disputed script with other examples of a person's writing to investigate their common authorship? If so, what writings could be used as exemplars? What alternative means besides such direct factfinder comparison might be attempted or required?

The answers to these questions were never totally clear. As to whether the factfinder would be allowed direct comparison, the answer seems to have been "yes", at least in civil cases.<sup>90</sup> This answer was qualified, however. In general, such factfinder direct comparison could not constitute the sole proof of authorship: one might say there was a fixed rule of insufficiency.<sup>91</sup> Instead the proponent was required to present a recognition witness—someone with a direct acquaintance with the asserted author's hand<sup>92</sup> (though we shall see presently just how shallow an acquaintance would suffice). This recognition witness also served as the conduit through which exemplars (if any were to be used) were authenticated and passed to the jury.<sup>93</sup> Just before 1800, important changes in this approach occurred. Previously, contested unwitnessed authorship of a document could be proved in a criminal case in the manner outlined above only when the document was found in the

---

<sup>89</sup> Wigmore, *Proof by Comparison of Handwriting; Its History*, 30 AM. L. REV. 481, 494 (1896).

<sup>90</sup> See *id.* at 486-93 (collecting other authorities). There seems to have been an exception even to the requirement of a recognition witness in the case of an ancient document, when direct comparison alone was allowed to be the sole source of information for the jury. See *id.* at 486 n.4.

<sup>91</sup> See *id.* at 487 n.3 (quoting F. BULLER, TRIALS AT NISI PRIUS 231 (1767): "[Comparison of hands] may be used as circumstantial and confirming evidence, if the Fact be otherwise proved.").

<sup>92</sup> See *id.* at 487 & n.3, 492-93.

<sup>93</sup> See *id.* at 488-89.

defendant's custody.<sup>94</sup> Otherwise the document was formally unprovable.<sup>95</sup> By 1800 this limitation had been abolished, and a single approach to the identification of denied handwriting was followed in both civil and criminal cases.<sup>96</sup> Further, the class of persons deemed acceptable as recognition witnesses was expanded to include correspondents and others who had never actually seen the putative author write.<sup>97</sup> While the doctrine of identification by a recognition witness was expanded, however, it also was subjected to one important constraint. Recognition witnesses were prohibited from showing the jury exemplars of the putative author's writing that were relevant only in that they showed the form of that writing. Instead, (presumably because of the perceived problems of "collateral" authentication) any exemplars had to be independently relevant to and admissible in the case. Documents whose relevance was solely that they represented examples of a putative author's normal handwriting were no longer allowed.<sup>98</sup>

While one form of information, exemplars, was restricted, the primary key to the authentication system, the recognition witness, remained breathtakingly free of constraint. Prior to 1800, direct comparison seems to have been viewed as a potential source of rational information for the jury. Nevertheless, direct comparison without some form of corroboration was considered dangerously ambiguous and indeterminate; it was still necessary to provide the formal minimum of proof when no exemplars of any kind were available. This formal minimum, which also served to corroborate the evidence of factfinder comparison when exemplars were available, could not be made too onerous, or the result would be too many inconvenient formal failures to prove documents. The compromise was a doctrine that strikes nearly everyone as ludicrous when first encountered but that survives to the present day.<sup>99</sup> Any person who has observed a person write even once, even years ago, may testify to recognizing handwriting, and that testimony is formally sufficient to take the issue to the jury.<sup>100</sup>

---

<sup>94</sup> See *id.* at 485.

<sup>95</sup> See *id.* at 485-86.

<sup>96</sup> *De la Motte's Case*, 21 How St. Tr. 803, 810-11 (1781).

<sup>97</sup> The only acceptable witnesses were those "who had seen the person write . . . [or] who had received writings subsequently treated by him as genuine or who had the custody of ancient documents of the same person's." *Id.* at 493. Such witnesses were permitted to "merely examin[e] the disputed writing and . . . bring[] into court the specimens they knew and juxtapos[e] them." *Id.* Wigmore concluded that "the classes of witnesses who may testify to handwriting have increased in numbers by successive enlargements." *Id.* at 499.

<sup>98</sup> See Wigmore, *supra* note 89, at 492-93.

<sup>99</sup> See C. McCORMICK, *supra* note 15, § 221, at 689-91.

<sup>100</sup> See *id.* at 690.



Let us see where that would leave us about 1810, after the general expansion of the doctrine to both civil and criminal cases and the restriction on the display of otherwise irrelevant exemplars. A clerk who had seen someone write only once, two decades earlier, could authenticate a challenged document. A challenger, however, could not point out the most blatant differences between the purported writing and hundreds of pages of undoubtedly authentic daily journals in the putative author's hand, because they were not independently relevant to the case. Given such possibilities, the new restriction on the display of exemplars seemed intuitively irrational. The next half-century saw a continuing skirmish over when not-otherwise-relevant documents could be brought in as exemplars and when any document could be used by the jury for direct comparison.<sup>101</sup> This esoteric quarrel occasioned an accidental conquest by a nonexistent expertise.

Surprisingly, the first recorded opinion addressing the propriety of anything resembling expert testimony in any context did not occur until Lord Mansfield admitted an engineer's testimony in *Folkes v. Chadd*<sup>102</sup> in 1782.<sup>103</sup> The first case on record in which a witness was proffered as specially qualified to testify solely from direct comparison between exemplars and a challenged writing rather than from "recognition" was *Goodtitle d. Revett v. Braham*.<sup>104</sup> There Lord Kenyon, on the authority of *Folkes v. Chadd*, admitted the testimony of two inspectors of franks, postal officials whose duties entailed checking the authenticity of the signatures of members of Parliament and other officials on mail dispatched pursuant to their franking privilege.<sup>105</sup> The next

<sup>101</sup> See 7 J. WIGMORE, *supra* note 15, § 1994.

<sup>102</sup> 3 Dougl. 157, 99 Eng. Rep. 589 (K.B. 1782).

<sup>103</sup> See *supra* note 7. This case was not the first use of expert testimony, of course, merely the first opinion to analyze it consciously.

<sup>104</sup> 4 Term Rep. 497 (1792). The claim has been made that such skilled witnesses were heard in the ecclesiastical courts long before this. See *Hanriot v. Sherwood*, 82 Va. 1, 13 (1884) (referring to the admission of testimony on the genuineness of writings or a signature: "[t]his sort of testimony was always admitted by the ecclesiastical courts"); 7 J. WIGMORE, *supra* note 15, § 1993, n.5; *On the Proof of Handwriting*, 2 L. REV. & Q.J. BRIT. & FOREIGN JURISPRUDENCE 285, 294 (1845) (unsigned article). Whether the sources cited refer to such witnesses, however, or to the judge's right as a literate person to make direct comparison aided by forensic argument by the parties, is anything but clear. What is clear is that by the early 19th century the ecclesiastical courts conformed to the common law rules, and their use of handwriting experts, if it ever existed, was "totally obsolete." *On the Proof of Handwriting*, 7 MONTHLY L. MAG. 120, 136 (1840) (unsigned article). What would have counted as a handwriting expert in ecclesiastical practice in the 17th century, if indeed such were allowed, seems totally beyond reconstruction.

<sup>105</sup> The case report contains an interesting summary of the examiners' approach.

On cross-examination, they admitted, that they had never detected an imitation of the hand of a very old person, who wrote with difficulty, and might be supposed frequently to stop. That their principal means of

year, in *Carey v. Pitt, Esq.*,<sup>106</sup> Kenyon refused to admit the identical kind of evidence from an inspector of franks, declining to treat the inspector differently from a lay witness. Kenyon acknowledged that such evidence had been received in *Revett v. Braham*, but held that its admission was insignificant in that "he had, in his charge to the Jury, laid no stress upon it."<sup>107</sup>

Later that same year, in *Stranger v. Searle*,<sup>108</sup> another litigant called an inspector of franks to testify that the signature of the defendant's name on a bill of exchange was in fact the defendant's. The witness had neither seen the defendant write nor corresponded with him and thus did not meet the accepted requirements for a recognition witness. Kenyon would not let him testify, refusing to expand the criteria of qualification for a recognition witness even though the circumstances indicated that the inspector had inspected numerous real franks signed by the defendant. Interestingly, Kenyon indicated that he might admit testimony by an expert claiming the ability to distinguish a feigned from a natural hand without using a comparison document, but the witness admitted he could not do so.

Finally, in *R. v. Cator*,<sup>109</sup> a criminal libel trial, the entire issue was argued at length before Baron Hotham. The prosecution sought to call an inspector of franks to testify that a) the writing on the libelous letters was a disguised hand and b) by comparing the disguised hand of the libels with exemplars of the natural hand of the defendant, he had determined that the defendant wrote the libels. The prosecution asserted that the inspector of the franks was "a man of science,"<sup>110</sup> without apparently providing a basis for that contention. The judge split his decision, though not quite down the middle. He decided that the inspector could testify as to whether he believed the libels to be in a disguised hand if he could do so from looking at only them, but that no one could testify to the identity of the author by comparison of hands. Interestingly, identifying a writing as a feigned hand by merely looking

---

knowing was by seeing whether the letters were painted, that is, gone over a second time with the pen, which they admitted might happen to any person from a failure of ink.

*Goodtitle*, 4 Term Rep. at 498. It remains unclear how this technique qualified them to say that the will in question was written by the deceased who wrote the exemplars.

<sup>106</sup> Peake Add. Cas. 130 (1793).

<sup>107</sup> *Id.* at 131.

<sup>108</sup> 1 Esp. 14 (1793).

<sup>109</sup> 4 Esp. 117 (C.P. 1802).

<sup>110</sup> *Id.* at 134. Similarly, both Evans and Bentham, in their early 19th century commentaries on English evidence, assume the existence of expertise without enlightening us as to the basis for their acts of faith. See J. BENTHAM, *supra* note 83, at 599; 2 W. EVANS, NOTES TO POTHIER 159 (No. 16 VI) (1806).

at it is precisely what the inspector of franks in *Stranger v. Searle* testified that he could not do. Perhaps the skill of postal inspectors had increased in the intervening six years.

Three aspects of the testimony offered in these cases, aspects that hold true in all the cases antedating the Common Law Procedure Act of 1854, demand notice. First, the expertise being asserted is of a fairly low order, not one embedded in a structure of mystifying theory, but rather one founded on a kind of honest mechanic's empiric that might be easier for a jury to evaluate rationally.<sup>111</sup> Second, it is the kind of testimony that all handwriting experts testifying today would condemn as unscientific and undependable;<sup>112</sup> the "experts" of past centuries almost certainly would not be allowed to testify today. Third, until the Act was passed, such testimony was generally inadmissible in England, although the whole issue was rehashed at length before the Kings Bench *en banc* in *Doe d. Mudd v. Suckermore*.<sup>113</sup> That 1836 case ended in a split opinion—two justices for exclusion and two for limited admission—which resulted in continued exclusion.<sup>114</sup>

Ultimately, the law was changed by the interpretation the courts placed on section twenty-seven of the Common Law Procedure Act of 1854. The Act provided:

Comparison of a disputed writing with any writing proved to the satisfaction of the judge to be genuine shall be permitted to be made by witnesses; and such writings, and the evidence of witnesses respecting the same, may be submitted to the Court and jury as evidence of the genuineness, or otherwise, of the writing in dispute.<sup>115</sup>

The terms of this statute do not on their face address the question of the reality of expertise. They could as easily, if not more easily, be read to resolve the controversy concerning extraneous exemplars, with the term "witness" interpreted to mean "recognition witness" as a conduit for placing these extraneous exemplars before the jury. Additionally, the broad term "witness" might mean asserted experts. This it was taken to do, though not immediately.<sup>116</sup> Eventually it seems to have

---

<sup>111</sup> See *supra* note 105.

<sup>112</sup> See, e.g., A. OSBORN, QUESTIONED DOCUMENTS 286 (2d ed. 1929) (decrying the unreliability of bank employees' conclusions).

<sup>113</sup> 5 A. & E. 703 (K.B. 1836).

<sup>114</sup> The *en banc* decision followed a trial court decision to exclude testimony based on direct examination of a witness who had never seen the party in question write. See *id.* at 705.

<sup>115</sup> Common Law Procedure Act, 1854, 17 & 18 Vict. ch. 125, § 27.

<sup>116</sup> For the application of the Act to non-expert authentication of exemplars solely for jury comparison, see *Cresswell v. Jackson*, 2 F. & F. 24 (1860); cf. *Doe d. Devine*

been interpreted as a statutory instruction to accept proffered "experts" without much examination of their methods or their proficiency. 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—and that largely as a result of the ingenuity of lawyers. It is very little exaggeration to say that Parliament, intentionally or not, mandated the admission of testimony by practitioners of this "expertise" even before it came into existence. This void was to be filled by foreign import.

In 1871 Charles Chabot published *The Handwriting of Junius, Professionally Investigated*<sup>117</sup> in England. This book was the first in English to claim that a science of handwriting identification existed and to explicate the claimed discipline. Almost simultaneously, the subject of handwriting identification leapt into the public consciousness because of its role in the most publicized legal affair of 19th century England, the Tichborne Claimant case.<sup>118</sup> Chabot was called by the prosecution to give testimony in the criminal proceeding of that controversy, thus giving the "discipline" its first of two major public relations coups.<sup>119</sup>

It is not surprising that the foundations of today's "science" of

---

v. Wilson, 10 Moo. 502 (Privy Council 1855). The first reported case construing the statutory language to authorize expert comparison and testimony appears to have been *Regina v. Harvey*, 11 Cox Crim. L. Cases 546 (Assizes 1869) (actually, *Harvey* construed an almost verbatim provision of the Criminal Procedure Act of 1865). However, the allowance of expert testimony seems to have become established in practice by 1864. See Jury charge of Cockburn, C.J. in *Creswell v. Jackson*, 4 F & F. 1, 8 (1864), (directing caution in evaluating such expert witnesses.) By 1869, the court in *Harvey*, seemed to indicate that direct juror comparison would not be allowed "without the assistance of an expert" 11 Cox Crim. L. Cases at 548.

<sup>117</sup> Actually, its publication seems to have been procured by M. Chabot's English disciple Edward Twistleton, who wrote the introduction. Osborn was a great admirer of Twistleton, and he referred to Twistleton's introduction to Chabot's book as "one of the best discussions of the principles of handwriting identification that has ever been issued." A. OSBORN, *supra* note 112, at 694.

<sup>118</sup> "The Tichborne Claimant case," really two cases, involved a man described by the *Sunday Times* as a "coarse and unlettered butcher," who in 1867 claimed that he was the long-lost heir to the Tichborne baronetcy and estates. The heir, Roger Tichborne, had been given up as drowned off the coast of South America in 1854; but when Arthur Orton stepped forward thirteen years later claiming to be Roger, Roger's mother reportedly believed the story. The rest of the family repudiated him as an imposter. See D. WOODRUFF, *THE TICHBORNE CLAIMANT*, at xiii-xiv (1957). After a civil trial, which determined that Orton was an imposter, see *Tichborne v. Lushington*, Notes of Proceedings 5212 (1872), he was tried as a criminal. See *Regina v. Castro*, 9 L.R.-Q.B. 350 (1874). It was at the latter proceeding that handwriting analysis was used, by both the prosecution and the defense. An old friend of Roger's (and believer in Orton's claim), Guilford Onslow, testified that Orton's handwriting and spelling were that of a gentleman. See D. WOODRUFF, *supra*, at 381.

<sup>119</sup> See D. WOODRUFF, *supra* note 118, at 283. Its second major coup was, of course, the Lindbergh Baby kidnapping case. See *infra* notes 175-178 and accompanying text.

handwriting analysis came out of 19th century Continental thought. Western Europe was then a hotbed of extreme rationalism<sup>120</sup> on the fringes of science, where theories were spun out to satisfyingly mystical complexity and experience was expected to conform or be damned. The same intellectual climate that gave us phrenology,<sup>121</sup> Lombrosian physiognomy<sup>122</sup> and, as previously noted, graphology,<sup>123</sup> gave us "chirography," or handwriting identification as a "science."<sup>124</sup>

Some critics quickly detected the fundamental weaknesses of the entire discipline. In a review of Chabot's book published in the London Times of May 22, 1871, A. Hayward, Esq., Q.C., examined the assumptions behind Chabot's methodology and found them largely unsupported:

[O]n almost every occasion when their evidence is adduced, it is adduced on both sides. The process is invariably the same. To prove identity, they pick out all the similar words, letters and marks: to disprove it, the dissimilar. . . .

---

<sup>120</sup> The term "rationalism" is used here in its classic philosophical sense: the belief that the most valid reality is represented by the constructs of the mind, more or less independent of external sense data. Thus, "Platonism" is a version of rationalism. Rationalist epistemology was classically opposed to that of "empiricism," which held that ultimate reality was external, knowable only imperfectly and only by careful examination of sense data. In the 18th century these two approaches to knowledge were in continual struggle. Modern attempts to join these strains of rational and empirical thought in the philosophy of science, plus popular use of the terms interchangeably to mean "reasonable based on good (usually empirical) evidence" have led to a blurring of these lines. *See generally* 7 *ENCYCLOPEDIA OF PHILOSOPHY* 69-74 (P. Edwards ed. 1967) (entry under "rationalism").

<sup>121</sup> "[T]he pseudo-science that relates the mental and temperamental characteristics of an individual to the bumps and hollows of his skull." 22 *ENCYCLOPEDIA AMERICANA* 22 (1976). First put forth by the German physician Gall and his disciple Spurzheim in the early 19th century, it was not finally rejected as unfounded until the second half of that century, and in the interim had many adherents and practitioners. *See id.* at 22-23.

<sup>122</sup> Italian criminologist Cesare Lombroso was responsible for the influential 19th century school of "scientific" physiognomy. He asserted that criminal types could be identified by physical "stigmata". Until finally exploded by Charles Goring in 1913, his theories had many adherents. *See* 7 *ENCYCLOPAEDIA BRITANNICA MICROPAEDIA* 980 (15th ed. 1974).

<sup>123</sup> *See supra* note 13.

<sup>124</sup> France continues to have her own troubles with this offspring, as is illustrated by the Villemin murder case. In October of 1984, four-year-old Gregory Villemin was kidnapped and a series of handwritten revenge notes were received. After his body was discovered, a witness incriminated the boy's uncle. After the uncle was identified by a handwriting expert as the author of the notes, he was indicted. Later, his attorney called for a more controlled examination by another expert. In a blind test involving exemplars from all the members of the family (which had apparently been supplied only for test purposes), Gregory's mother was identified as the author, and she was indicted. *See* Int'l Herald Tribune, July 18, 1985, at 2, col. 3. Thus two defendants were indicted as the murderer on the basis of being identified by two different "experts" as the sole authors of the notes.

....

....

If, as we are told, most feigned hands are formed on the same principle—namely, by reversing the slope of the downstrokes and making the round turns of the letters angular, or *vice versa*, most feigned hands must have a general resemblance, and the feigned hand of one person may easily be confounded with the feigned hand of another.

Listening to the evidence or reading the reports of experts, we often feel very much 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. After applying a succession of magnifying glasses we cannot make out more than two or three letters, or parts of letters, in the obliterated dates which Mr. Chabot is able to reproduce in full; and in each instance the day of the month is to our eyes a blackened spot. His method must be that of the comparative anatomist, like Cuvier or Owen, who can arrive at the form and habits of an extinct animal by induction from a bone; and, after a vain attempt to follow him, we recognize the justice of Mr. Twistleton's remark that "it is not the merely having bodily eyes which enables any one to see in two documents either discrepancies or coincidences." We lack the spiritual sense, or we lack faith.<sup>125</sup>

Prior to the passage of the English statute, most American courts excluded proffered "expert" testimony, following the English practice.<sup>126</sup> There were, however, a few interesting exceptions. In 1836, in an opinion distinguished for its opacity,<sup>127</sup> Massachusetts did what England did not do in *Doe d. Mudd v. Suckermore*.<sup>128</sup> It accepted expert testimony based on the comparison of challenged documents with exemplars. Massachusetts accepted the rationale rejected by the court in *Suckermore*: that such testimony could not be of lower quality than the recognition testimony usually accepted.<sup>129</sup> The Massachusetts deci-

---

<sup>125</sup> The Times (London), May 22, 1871, at 8, col. 1. (letter titled "The Handwriting of Junius").

<sup>126</sup> Data on acceptance by each American jurisdiction can be found in Appendix 3, *infra*, in which they are arranged chronologically.

<sup>127</sup> See *Moody v. Rowell*, 34 Mass. (17 Pick.) 490 (1836).

<sup>128</sup> 5 A & E 703 (K.B. 1836).

<sup>129</sup> The opinion was hardly a wholehearted endorsement of expertise, as the court concluded that "this species of evidence, though generally very slight, and often wholly immaterial, is competent evidence." *Moody*, 34 Mass. (17 Pick.) at 498.

sion was followed in Ohio in 1850,<sup>130</sup> Maine in 1851,<sup>131</sup> and New Hampshire in 1852.<sup>132</sup> Even more interesting was an 1851 Iowa statute, the first in a common law<sup>133</sup> jurisdiction, explicitly providing for the admission of "expert" testimony.<sup>134</sup> Thus, when the English statute was passed, five American common-law jurisdictions already accepted such testimony while the other twenty common-law states then in the union rejected it.<sup>135</sup> By 1880, acceptance of expert testimony had been broadened by judicial decisions and legislation. Nevertheless, it remained inadmissible in a majority of American jurisdictions.<sup>136</sup> It was not until the American offspring of M. Chabot emerged in the last quarter of the 19th century that such testimony gained general acceptance, and not until well into the 20th century when the last state finally

---

<sup>130</sup> See *Hicks v. Person*, 19 Ohio 426, 441-42 (1850).

<sup>131</sup> See *Sweetser v. Lowell*, 33 Me. 446, 450 (1851).

<sup>132</sup> See *Bowman v. Sanborn*, 25 N.H. 87, 112 (1852).

<sup>133</sup> The phrase "common law" is important. Louisiana adopted a version of the Code Napoleon in 1808. That code provided for the use of skilled handwriting witnesses in cases involving questioned documents. See LA. CIV. CODE art. 226 (1808), reproduced *infra* Appendix 3. This early adoption is consistent with the French origin of the "expertise," see *supra* notes 117 & 124 and accompanying text, and may perhaps indicate the existence and influence of a cadre of self-styled "scientific" experts in France around the turn of the 18th century. Be that as it may, there is no reported decision indicating that this section of the Louisiana code was used. The few reported decisions deal with recognition witnesses, though the code section is noted. See *Plicque v. Labranche*, 9 La. 559, 562 (1836); *Sauve v. Dawson*, 2 Mart. 202 (1812). Therefore, the real situation in Louisiana is uncertain.

<sup>134</sup> See IOWA CODE § 2404 (1850-51). Iowa adopted the Field Code in 1850-51, and it is fairly clear that the source of the handwriting expert statute was a similar provision, section 4 of the 1848 Field Code, a set of evidence code provisions never adopted in New York. See Final Report of the Commissioners on Practice and Procedure, Part IV, § 1764 (1848). The New York version did not use the word "expert," an adjective just coming into use as a noun, but the anonymous redrafter of the Iowa version adopted the neologism. Given Jeremy Bentham's attitude toward experts in general and the possibility of handwriting experts in particular (which he referred to as "scientific witnesses," see *supra* note 110) it is not surprising that this first statute was the product of the American Benthamite movement that produced the drafting and widespread adoption of the Field Code.

<sup>135</sup> See *infra* Appendix 3.

<sup>136</sup> See *infra* Appendix 3; see also Lawson, *Proof of Handwriting: II—By Comparison*, 17 AM. L. REV. 21, 31-60 (1883) (collecting data). Up to 1900, the bulk of the acceptances were by decision; after 1900, by statute. Interestingly, from 1925 to 1968 no additional jurisdiction formally accepted handwriting expert testimony, and three jurisdictions produced no formal manifestation of acceptance until the 1970s. Indeed, in Wyoming there is still no formal authority. As late as 1925, Arkansas was still getting along nicely without experts, using only circumstantial evidence, direct jury comparison, and forensic argument by counsel.

Even in the five jurisdictions, however, that throughout the middle third of this century lacked formal authority for acceptance, we believe that handwriting expertise was probably accepted in practice after *State v. Hauptmann*, 115 N.J.L. 412 (1935). For a discussion of the role played by handwriting expertise in that notorious case, see *infra* notes 174-77 and accompanying text.

abandoned the common law position.

### III. THE PERIOD OF TRANSITION: SIMULTANEOUS JUDICIAL ACCEPTANCE OF AND HOSTILITY TOWARD HANDWRITING EXPERTISE.

Between 1852<sup>137</sup> and the beginning of the Civil War, no additional American jurisdiction accepted testimony by putative handwriting experts. Between 1861 and 1883, however, ten more jurisdictions accepted such testimony by case decision while seven more did so by statute. Thus, by 1883, twenty-two common-law jurisdictions admitted it, at least under certain conditions, and seventeen did not.<sup>138</sup> In 1900 a number of states still held out against admitting such testimony.<sup>139</sup> A federal statute mandating admission cured the split among federal courts in 1913, and by 1925 all but five jurisdictions had declared handwriting expertise admissible.<sup>140</sup>

Two aspects of these developments are important for our purposes. First, though the "moving spirit of the age" seemed to favor allowance of "expert" handwriting testimony and judicial decisions were the source of change in most jurisdictions, a strong, simultaneous counter-current of judicial skepticism and hostility existed toward that expertise. This skepticism was sometimes manifested even by some of the same courts that established admissibility in the first cases. To understand this seeming contradiction, one must remember that the dominant rationale for allowing such testimony in the antebellum decisions,<sup>141</sup> and even some post Civil War decisions,<sup>142</sup> was simply that experts could not possibly be worse than recognition witnesses. The skepticism of some late 19th and early 20th century courts was even greater. Consider the following quotations from New Jersey, in which admissibility resulted by statute: "All doubts respecting the competency of the opinion of experts in handwriting based upon mere comparison, as evidence, have been removed by statute; but it still must be esteemed proof of low degree."<sup>143</sup> Also consider the opinion of the New York Court of

---

<sup>137</sup> See *supra* note 132 and accompanying text.

<sup>138</sup> See *infra* Appendix 3; see also Lawson, *supra* note 136, at 30-60 (analyzing the response of each state to the handwriting experts).

<sup>139</sup> See *infra* Appendix 3.

<sup>140</sup> See *id.*

<sup>141</sup> See *supra* notes 127-32 and accompanying text.

<sup>142</sup> See, e.g., *Miles v. Loomis*, 75 N.Y. 288, 295 (1878); see also *People v. Collins*, 94 Misc. 2d 704, 720, 405 N.Y.S.2d 365, 375 (Sup. Ct. 1978) (asserting that this rationale, astoundingly, continues to be the justification for admissibility).

<sup>143</sup> *Mutual Benefit Life Ins. Co. v. Brown*, 30 N.J. Eq. 193, 201 (1878) (citation to statute omitted); see *infra* Appendix 3.



Appeals in *Hoag v. Wright*, in which admissibility derived from an earlier decision of the same court:

The opinions of experts upon handwriting, who testify from comparison only, are regarded by the courts as of uncertain value, because in so many cases where such evidence is received witnesses of equal honesty, intelligence and experience reach conclusions not only diametrically opposite, but always in favor of the party who called them.<sup>144</sup>

Courts' hostility to the unwelcomed force of change may account for another characteristic of their decisions in the late 19th century, one that later had curiously ironic results. The typical judicial decision allowing in expert handwriting testimony opened the door to it as narrowly as possible. Specifically, it carried forward the common law restriction on "extraneous" exemplars and required that the experts ply their trade only on independently relevant documents.<sup>145</sup> While that restriction on the *experts* seems to have eroded over time through the

---

<sup>144</sup> *Hoag v. Wright*, 174 N.Y. 36, 42, 66 N.E. 579, 581 (1903); see also *In re Fuller's Estate*, 222 Pa. 182, 183, 70 A. 1005, 1006 (1908) (finding that such testimony is of "very little" reliability). Scotland also remained skeptical. An unsigned article entitled *Comparatio Literarum* in the *Scottish Journal of Jurisprudence*, reprinted in 12 ALB. L.J. 315 (1875), observed:

[I]n Scotland we have a great distrust of mere experts [on handwriting]—and justly. Like mad doctors, they support the side on which they are engaged. Also, they found too much on very minute circumstances of similarity or dissimilarity. . . .

[One] court, while declining actually to hold that such evidence was incompetent, stated that it would not have the slightest effect on their minds, and it was for the parties to consider whether it was worth while to incur useless expense. . . . On this matter, Lord Justice Clerk (Lord Moncreiff) made these remarks:

"... During the debate something was said about experts. I think that a person acquainted with the particular handwriting of the party, or even an ordinary person unacquainted with it, is as likely to ascertain the truth of the matter as those witnesses who, looking at it from a scientific point of view, are apt frequently to exaggerate small distinctions. Such witnesses are in a different position from men of skill in such matters as chemical and mechanical facts, which are not ordinary observers."

... Lord Gifford remarked: "The evidence as to handwriting is unsatisfactory. I do not attach much importance to the absence of the evidence of experts, for such evidence is more in the nature of argument than of a proof. All they can do is merely to point out resemblances or differences in the disputed signature as compared with genuine specimens of handwriting."

*Id.* at 316.

<sup>145</sup> See virtually any 19th century decision in Appendix 3. See *Tyler v. Todd* 36 Conn. 218, 223 (1869); *Rogers v. Tyley*, 144 Ill. 652, 665, 32 N.E. 393, 397 (1892); *State v. Scott*, 45 Mo. 302, 305-06 (1870).

influence of Osborn and Wigmore,<sup>146</sup> it was often irrationally retained in a way that prevented the *opponent* of the expert from putting forth rationally fair and useful tests of the expert on cross examination. We shall return to this subject below. First we must examine how we moved from the underlying judicial hostility of the late 19th century to the generally uncritical acceptance of today.

#### IV. THE CRUSADE FOR RESPECT FOR HANDWRITING EXPERTISE

The elimination of the judicial attitude of contempt and the treatment of handwriting identification expertise by rules appropriate to a valid scientific discipline became a thirty-year crusade for an unlikely duo: John H. Wigmore<sup>147</sup> and Albert Osborn.<sup>148</sup> Wigmore and Osborn came from backgrounds that were both greatly similar and greatly dissimilar. Osborn was born in 1858 and Wigmore in 1863. Both were white middle-class males, but they grew up in different places and toward opposite ends of such middle-class social scale as existed in their communities. Wigmore's father was a well-off and educated merchant in San Francisco, and Wigmore himself went east to receive the ultimate in establishment formal academic education: Harvard College, B.A. 1883; Harvard M.A and LL.B., 1887.<sup>149</sup> Osborn's parents were working farmers in Grass Lake, Michigan, and his formal education was incomplete.<sup>150</sup> During his one-year residency at Michigan State College, Osborn became fascinated with penmanship,<sup>151</sup> the skill of being able to write fluidly in near-perfect copybook styles. One must remember that the typewriter had yet to be perfected in a commercially usable form; the copying of documents was still done entirely by an army of penmen. Mastery of penmanship offered a path off the farm within the scope of both Osborn's means and his horizons. Apparently he had the knack, for he became a master ambidextrous penman, largely self-taught. He also had remarkable ambition. He did not leave

---

<sup>146</sup> See *infra* notes 147-181 and accompanying text (recounting Osborn's and Wigmore's crusade on behalf of handwriting expertise).

<sup>147</sup> The details of Wigmore's life given here are drawn from W. ROALFE, JOHN HENRY WIGMORE (1977). The interpretations are the authors' own. For a similar interpretation of Wigmore's views and personality, see Graham, "There'll Always be an England": *The Instrumental Ideology of Evidence*, 85 MICH. L. REV. 1204, 1216-27 (1987).

<sup>148</sup> Details of Albert Osborn's life are taken from a laudatory sketch by Clark Sellers, *Albert Sherman Osborn: Questioned Document Pioneer*, 45 A.B.A. J. 1285 (1959). The interpretations are the authors' own.

<sup>149</sup> See W. ROALFE, *supra* note 147, at 9-10.

<sup>150</sup> Osborn attended Michigan State in Lansing, but an eye ailment forced him to return home after one year. See Sellers, *supra* note 148, at 1286.

<sup>151</sup> See *id.*

the farm to become a copyist but stayed there and honed his skill in his spare time until a better opportunity arose. He became a teacher of penmanship at an established business college in Rochester, New York.

Osborn left the farm for the Rochester Business Institute in 1882, four years after the first sales of the mass-produced upper and lower case Remington typewriter<sup>152</sup> that, along with imitations and newer models, was eventually to put penmen-scribes out of business. After a few years as teacher of penmanship, the bright young Osborn must have seen that he was in a dying profession, as low-paid women of six weeks training began producing more letters in a day than the experienced penmen with extensive training whom they replaced. Young Albert had already spotted an option, however, and became intensely interested in it (intensity was one of his dominant character traits). There was an attractive new market for "expert" opinions on handwriting identification as the American courts opened up to such testimony, and no real experts existed to fill the demand. Under the direct influence, we may hypothesize, of Charles Chabot's book, Albert Osborn set out sometime in the 1880s to become a founding father of "scientific" handwriting identification in the United States.<sup>153</sup>

There are more than a few ironies in what can be gleaned of Osborn's personality and attitudes from his rather voluminous writings over the next four decades.<sup>154</sup> First, he was clearly a man of exceptional intelligence. Whatever errors he fathered and bequeathed, they were not the errors of a stupid person, or even a person of average critical abilities. Second, he viewed himself as a kind of reforming iconoclast in the mold (whether he was directly aware of the mold is another matter) of Jeremy Bentham. Third, he had a faith in the ability of the human mind to create a practical expertise for the solution of any class of

---

<sup>152</sup> See THE NEW COLUMBIA ENCYCLOPEDIA 2812 (4th ed. 1975).

<sup>153</sup> Osborn has two rivals for the founding fatherhood: William E. Hagan, who published *Disputed Handwriting* (1894), and Daniel T. Ames, who published *Ames on Forgery* (1899). Osborn acknowledges their contributions in his critical bibliography in the 1929 second edition of *Questioned Documents*, if nowhere else. The reason his name lives on and their's sank into obscurity seems largely due to Osborn's force of personality, his longevity, and his relationship with Wigmore. On Osborn's admiration for Twistleton's introduction to Chabot's book, see *supra* note 117.

<sup>154</sup> Osborn published two editions of *Questioned Documents* (1910 and 1929), the latter of which ran to 1042 pages. He also published *The Problem of Proof* (1922), a general work on inference, proof and trial practice for which Wigmore wrote a glowing introduction, and *Questioned Document Problems* (1944) (with his son, Albert D. Osborn). The latter was dedicated to the recently deceased Wigmore. Finally, he wrote a number of articles, such as *Handwriting Expert Testimony*, 65 ALB. L.J. 9 (1903) and *Progress in Proof of Handwriting and Documents*, 24 J. CRIM. L. & CRIMINOLOGY 118 (1933) (addressing progress not in handwriting identification techniques but in the acceptance of the area of expertise by the courts).

problems and a curious compartmentalization of the empirical and rational bases for his own beliefs. For instance, he invented the comparison microscope to obtain more and better empirical data upon which to apply his theories<sup>155</sup> but never subjected his theoretical postulates to any kind of organized examination and critique. These characteristics—faith in expertism and an unexamined rationalist belief in coherent and pretty but untested theories—were even more common in the late 19th and early 20th centuries than they are today. Osborn apparently joined them with vanity, arrogance, force of personality, and a rhetorical knack for making both his unexamined theories and his conclusions sound like empirically verified scientific certainties. This personality package was not at all uncharacteristic of successful men in and around the legal profession and elsewhere at the time. Roscoe Pound springs to mind. So, more to the point, does John H. Wigmore.

Wigmore attended Harvard early in the reign of Dean Langdell and was exposed to the excitement of the new the case method experiment. One can view Langdell's attitude toward the law at many different interesting levels of magnification, but one component of that attitude seems certainly to have been a reverence for and desire to capture the mantle of scientific expertism.<sup>156</sup> It had not been a century since the term "science" had ceased to be commonly applied to any organized critical examination of a subject matter and started to be reserved for things that were at least potentially subject to the method of empirical verification.<sup>157</sup> The status of the "scientist" has never been higher, as advances in both the theoretical and applied sciences presented the society with almost daily wonders. To borrow forms from the natural sciences in ways that were generally no more than rough metaphors would allow other disciplines (like law) to assert for themselves the status of "sciences" in their own right. It was in part this scientific mystique that led Langdell and his followers to espouse a "scientific" or "inductive" method of legal education.<sup>158</sup> One should note that when one's claim to being a practitioner of "juridical science" rests upon such shaky ground, it does not behoove one to embrace very stringent standards for evaluating other people's claims to be "scientists." The dominant spirit of an entire era of legal thought was marked by a reverence for and faith in "scientific expertise" and a kind of gentleman's agree-

---

<sup>155</sup> See Sellers, *supra* note 148, at 1287.

<sup>156</sup> See Note, *The Scientific Model in Law*, 75 GEO. L.J. 1967, 1975-76 (1987).

<sup>157</sup> The use of the word "science" to describe subjects of study was the common 18th century usage. Examples include "historical, mathematical, . . . [and] moral" science. 9 THE OXFORD ENGLISH DICTIONARY 221 (1961).

<sup>158</sup> For a vicious (but justified) critique of Wigmore's commitment to inductivism, see Graham, *supra* note 147, at 1219-20.

ment not to question the claims of expertise of anyone who showed up at the courthouse door not appearing visibly deranged and not practicing an already totally discredited discipline (such as phrenology or astrology), as long as these same people were willing to concede the primacy of "juridical science" within the legal sphere. Wigmore, like Pound, and a host of others, was a child of this era.

While at Harvard, Wigmore also came under the tutelage of James Bradley Thayer and, as a result, began his long love affair with the law of proof.<sup>159</sup> After graduating in 1887, Wigmore practiced law for two years but apparently didn't love the actual practice as much.<sup>160</sup> To break into teaching, the newly married Wigmore did something quite out of the ordinary. He accepted a job at Keio University in Japan to teach the Japanese the contours of Western, and in particular American, law.<sup>161</sup>

One of the striking things about Wigmore was his immense capacity for single-minded and concentrated work. During his three years in Japan he mastered the written Japanese language<sup>162</sup> sufficiently to return to the United States with the makings of his first multi-volume set of writing, treating Japanese law during the Tokugawa period.<sup>163</sup> Upon his return, he joined the faculty of Northwestern University Law School, where he was to remain for the rest of his life.<sup>164</sup> While law teachers at that time were expected to be generalists, as Wigmore certainly was, he quickly became heavily involved in his first love, the law of evidence, as a contributor of articles to journals and, in 1899, as one editor of the treatise authored by Simon Greenleaf.<sup>165</sup> When Thayer died without having produced the anticipated expansion of his *Preliminary Treatise*,<sup>166</sup> Wigmore stepped in as his natural successor and in 1904 came out with the four-volume first edition of his own, now-fa-

---

<sup>159</sup> Wigmore developed a deep respect for Thayer and felt indebted to him. Consequently, he made a special dedication "[t]o the memory of the public service and private friendship of two masters of the law of evidence, Charles Doe . . . and James Bradley Thayer of Massachusetts historian and teacher." See W. ROALFE, *supra* note 147, at 14 & n\*.

<sup>160</sup> It is odd that neither of the great chams of 19th and 20th century evidence theory, Bentham and Wigmore, ever tried a case. See Graham, *supra* note 147, at 1216.

<sup>161</sup> See W. ROALFE, *supra* note 147, at 21-31.

<sup>162</sup> See *id.* at 23.

<sup>163</sup> See 1 J. WIGMORE, LAW AND JUSTICE IN TOKUGAWA JAPAN at xi (1969).

<sup>164</sup> See W. ROALFE, *supra* note 147, at 32.

<sup>165</sup> See *id.* at 42 & n.78 (noting that Wigmore edited 1 S. GREENLEAF, A TREATISE ON THE LAW OF EVIDENCE (1899)).

<sup>166</sup> J. THAYER, A PRELIMINARY TREATISE ON EVIDENCE AT THE COMMON LAW (1898).

mous treatise.<sup>167</sup> From that point until his death in 1943 Wigmore was the 800 pound gorilla of American evidence law.<sup>168</sup>

If the whole law of proof was Wigmore's seigniory, the problems of proof surrounding challenged handwriting inhabited one fairly obscure corner of his domain. That corner nonetheless drew a substantial amount of his attention and interest throughout his career, both for its own sake and for the lessons he thought could be drawn from judicial treatment of the various types of evidence proffered there. In 1896 he published a fairly lengthy article in the *American Law Review* looking at the history of the subject although not specifically dealing with the validity of the expertise then gaining admittance to a majority of American courts.<sup>169</sup> Then, apparently sometime before 1910, he became friendly with Albert Osborn.<sup>170</sup>

Wigmore was an amateur of forensic science with great hopes for its assistance in the search for truth.<sup>171</sup> Though not incapable of a burst of energetic and telling skepticism, he seems to have been willing to accept many claims uncritically and at face value.<sup>172</sup> This certainly seems to have been the case with Osborn's "scientific" approach to handwriting identification. He advocated its acceptance by all courts in his introduction to the first edition of Osborn's book in 1910<sup>173</sup> and

---

<sup>167</sup> See W. ROALFE, *supra* note 147, at 77.

<sup>168</sup> Q: "Where does an 800 pound gorilla sit?" A: "Anywhere he wants to."

<sup>169</sup> See Wigmore, *supra* note 89, at 481-99.

<sup>170</sup> Wigmore's 1896 article, *supra* note 89, did not name particular handwriting experts or sources on such expertise. Nor were they mentioned in the sections on handwriting in his 1904 treatise, J. WIGMORE, A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE AT TRIALS AT COMMON LAW (3d ed. 1938), that seems to have been drawn largely from that article, *supra* note 89. Wigmore, however, wrote a long and admiring introduction for the first edition of Osborn's *Questioned Documents*, published in 1910. See *infra* note 173 and accompanying text.

<sup>171</sup> See W. ROALFE, *supra* note 147, at 61, 85-87 & 210; Graham, *supra* note 147, at 1223-24. Wigmore's interest in forensic science, sadly, became a literally life-long one. He died in the crash of a taxi he took after attending a meeting of the *Journal of Criminal Law and Criminology*. See W. ROALFE, *supra* note 147, at 275.

<sup>172</sup> One of the only printed attacks Wigmore ever launched against a "scientific" claim, see Wigmore, *Professor Muensterberg and the Psychology of Testimony*, 3 ILL. L. REV. 399 (1909), was triggered when the German psychologist Hugo Muensterberg overstepped the bounds of the gentleman's agreement and criticized the legal factfinding process as fundamentally irrational. See W. TWINING, THEORIES OF EVIDENCE: BENTHAM AND WIGMORE 135-36 (1985). The effectiveness of Wigmore's counterattack (even though contained in the somewhat juvenile format of a mock trial with characters such as R.E. Search and X. Perry Ment) should not obscure the point that Wigmore never examined the properly respectful "scientific" claims of those like Osborn. For more detail on Wigmore's scientism, set out in attempted sympathy, see W. TWINING, *supra*, at 135-42.

<sup>173</sup> In his introduction, Wigmore said of the field in general and Osborn's book in particular:

Well down into the 1800s, the most daring impositions remain possible.

continued to do so faithfully for the next thirty years. Together their advocacy brought "scientific" handwriting identification from a phenomenon barely tolerated by courts to a recognized source of useful and dependable information, despite the absence of a shred of empirical evidence of anyone's ability to do what Osborn claimed he and others of his trainees and followers could do.

The reader may judge how far Wigmore carried his part of the advocacy from the following passage from the third edition of his treatise:

It ought to be now well understood that the identification and the decipherment of documents, including handwriting and all other features, are the subject of scientific study, made by the aid of instruments of precision, and that modern research has elevated the whole subject into the realm of an applied science. Under such conditions, tangible reasons can be given for every opinion, precisely as an engineer can explain the reasons for constructing a bridge of a certain type. A qualified expert's opinion may therefore now be tested and judged by the reasons on which it is based. Such an opinion is as much more valuable than the untrained layman's opinion as that of a scientific and experienced geologist surpasses that of the empiric oil prospector.<sup>174</sup>

---

But society at last seems to have overtaken the falsifier once more. Science and art, in the mass, are more than a match for the isolated individual.

Throughout this book may be seen the spirit of candid reasoning and firm insistence on the use of it. . . . If judges and lawyers can thoroughly grasp the author's faith in the value of explicit, rational data for expert opinions, the whole atmosphere of such inquiries will become more healthy. The status of the expert will be properly strengthened . . . .

The reader arises with a profound respect for the dignity of the science and the multifarious dexterity of the art.

Wigmore, *Preface* to A. OSBORN, *QUESTIONED DOCUMENTS* at viii-ix (1910).

<sup>174</sup> 7 J. WIGMORE, *supra* note 170, § 1998, at 190 (citation to two of Osborn's books omitted). Anyone who still doubts the depths of Wigmore's commitment need only examine what he wrote concerning the few remaining examples of judicial skepticism.

He characterized the statement in *Toll v. Monitor B. & P. Co.*, 26 F.2d 51, 53 (8th Cir. 1928), that "evidence of experts based on comparison of handwriting is unsatisfactory and not very reliable" as follows: "[It] in this day and age is a statement so lacking in correct information as to be unworthy of a Federal court." 7 J. WIGMORE, *supra* note 170, § 2008, at 201 n.1.

And in criticism of a 1929 Michigan decision restricting expert testimony on the "ultimate issue" of genuineness, Wigmore declared:

[T]his astonishing rule is fortified by a quotation from an opinion rendered 50 years before modern science had supplied its special apparatus for accuracy, and is accompanied by an allusion to handwriting experts as

The capstone of all this triumph must be the trial of Bruno Hauptmann for the murder of the Lindbergh baby in 1936.<sup>175</sup> Osborn and a number of his followers<sup>176</sup> testified against Hauptmann, asserting the conclusion that Hauptmann had written the ransom note found in the baby's crib and later notes sent to the Lindbergh family. The bulk of Osborn's testimony could be the subject of the comment by Mr. Hayward on Mr. Chabot's book quoted above.<sup>177</sup> Perhaps Hauptmann

---

"usually employed and paid to express desired opinions" which is a gratuitous offense to the honest and competent practitioners of that science.

*Id.* § 2014 at 209 n.1 (criticizing *Morris v. Harris*, 247 Mich. 690, 226 N.W. 661 (1929)).

<sup>175</sup> See Sellers, *supra* note 148, at 1335.

<sup>176</sup> Osborn was followed on to the stand by seven more handwriting "experts" who claimed to have confirmed his identification by their own examinations: Elbridge W. Stein, John F. Tyrell, Herbert J. Walter, Harry M. Cassidy, William T. Souder, Clark Sellers, and Osborn's son, Albert D. Osborn. The cross examinations of these witnesses were performed ineffectively, and the defense later found that other experts refused to testify against Osborn and company. See J. FISHER, *THE LINDBERGH CASE* 341-47 (1987). This is consistent with inter-expert agreement on authorship or with inter-expert agreement on what was politically best for their field. As an aside, Russell Osborn, Osborn's grandson and Albert D.'s son, was later in the news as the expert who declared Clifford Irving's forgery of Howard Hughes' signature to be genuine. See S. FAY, L. CHESTER & M. LINKLATER, *supra* note 80, at 129-33.

<sup>177</sup> See *supra* text accompanying note 125. A substantial number of the "identifying" characteristics testified to by Osborn are not "chirographic" similarities but spelling errors. See Transcript at 950-1022, *State v. Hauptmann*, 115 N.J.L. 412, 180 A. 809 (1935). Hauptmann claimed he was made to copy the exemplars directly from the notes, spelling errors and all. Even assuming he did not, consider the following:

Since misspellings are the most obvious characteristics of a document to the non-technical eye, it would be an asset of the greatest possible value to examine and compare misspellings if we knew that in so doing we stood upon unassailable ground. But do we? Lest I be taken as a prejudiced party I will commence by again quoting Dr. Hans Gross, Professor of Criminology at Prague University, and author of the internationally famous classic on *Criminal Investigation* (citation omitted):—

"As to the spelling, upon which too much value is often placed, it proves absolutely nothing. . ."

" . . . Let us suppose that a native of Galway is brought before a higher Court in New York or Boston on the charge of writing a threatening letter, in which such words as "state" and "slow" are misspelled "shtate" and "shlow." In his specimen of handwriting the same words are similarly misspelt.

"He is almost certain to be convicted even in the absence of other evidence of weight, if his counsel is not sufficiently posted upon the Connaught dialect to know that almost every ill-educated Connaughtman in Boston or New York will do likewise."

An interesting example recently came before the writer's notice. A certain suspect was charged with the writing of a number of documents in which a misspelling of the words "at all" occurred, this being written as "atal". In Court the suspect was asked to write a test piece containing these words. He wrote them as "atal." What did this prove?

It established the fact that both the documents complained of and the text written in Court were probably the work of West-of-Ireland



was guilty, and perhaps he would have been convicted absent Osborn's testimony, but from a public perspective there seems to have been something else at work. People *wanted* Hauptmann to be guilty. They *wanted* to be sure of his guilt, and they were grateful to those who provided the forms of certainty regardless of rationality. Osborn became a kind of public hero, and the validity of "scientific" handwriting identification when practiced by a "competent" practitioner became public knowledge, which, like much public knowledge, turns out to be not so clearly right. This public anointment of handwriting expertise, however, (coupled with its judicial canonization in *Hauptmann*)<sup>178</sup> seems to have stamped out virtually all manifestations of judicial skepticism. The current status can best be illustrated by the following quote from *In re Estate of Sylvestri*,<sup>179</sup> a New York decision in sharp contrast to the thundered skepticism of *Hoag v. Wright*.<sup>180</sup> "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."<sup>181</sup>

#### V. CHOICES FOR THE FUTURE: EXCLUSION AND OTHER ALTERNATIVES

What should a lawyer do when confronted with the likelihood that a handwriting expert will be called by the opponent? First, a direct attack by motion in limine on such testimony ought to be successful, under either the so-called *Frye* test<sup>182</sup> or any other defensible standard.

---

men—nothing more and nothing less.

A. QUIRKE, FORGED, ANONYMOUS, AND SUSPECT DOCUMENTS 239, 241 (1930).

In the *Hauptmann* case, many of the misspellings may have been common to German immigrants of similar background and education, such as Isidore Fisch, to whom Hauptmann always attributed ownership of the ransom money found in Hauptmann's garage.

<sup>178</sup> 115 N.J.L. 412, 180 A. 809 (1935).

<sup>179</sup> 55 A.D.2d 916, 390 N.Y.S.2d 598 (App. Div. 1977).

<sup>180</sup> 174 N.Y. 36, 42, 66 N.E. 579, 581 (1903).

<sup>181</sup> *Sylvestri*, 55 A.D.2d at 918, 390 N.Y.S.2d at 600. The cynical observation refers to the *Hoag* court's belief that handwriting experts always reach conclusions in favor of the party who hired them. See *supra* note 144 and accompanying text. See also *Morrone v. Morrone*, 44 N.J. Super. 305, 312, 130 A.2d 396, 400 (App. Div. 1957) (noting that after *Hauptmann*, the testimony of handwriting experts could no longer be regarded as "the lowest order of evidence and . . . accorded little evidential weight.').

<sup>182</sup> See *United States v. Frye*, 293 F. 1013 (D.C. Cir. 1923). The *Frye* case is in reality an icon in the law of evidence that can be given many meanings. See generally McCormick, *Scientific Evidence: Defining a New Approach to Admissibility*, 67 IOWA L. REV. 879 (1982) (discussing the various approaches courts and commentators have taken in analyzing the *Frye* test). The case itself was an unlikely candidate for icon status. It is a two page opinion rejecting a *defendant's* proffer of exculpatory lie detector evidence, and the cynical might think the result could have been otherwise if the

The power of the argument is obvious. No cases have ever examined, much less determined, whether these "experts" can do what they claim. Further, the tests that have been done do not support their claim. 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.<sup>183</sup> Admission of evidence of no proven reliability as if it were at least minimally reliable denies the opponent a rational trial; therefore it denies the process due under our official assumptions about the function of trials, and it would appear to be a denial of due process even in a civil, much less a criminal, trial.

---

prosecution had offered the evidence. Be that as it may, the court in its operative paragraph opined that a new scientific process requires "general acceptance in the particular field in which it belongs" in order to be a proper subject for testimony. *Frye*, 293 F. at 1014. From this phrase has grown a forest of opinions and articles. See McCormick, *supra*, at 879-916. Most jurisdictions seem to have at one time or other claimed to be following the *Frye* test, just as today some claim to reject it, but what it is, exactly, is still subject to debate. See McCormick, *supra*, at 883-902; see also Gianelli, *The Admissibility of Novel Scientific Evidence: Frye v. United States a Half Century Later*, 80 COLUM. L. REV. 1197, 1228-31 (1980) (stating that the "current status of the *Frye* test is difficult to assess. . . . *Frye* may be tottering, but has not yet fallen."). At its broadest (and most defensible), it would stand for the proposition that all expert testimony requires some index of validity independent of the assertion of its practitioners and that this validation ought to meet some acceptable empirical standard. This is the sense in which we have used it. It might also be interpreted more narrowly, however, to apply only to "novel" evidence (grandfathering in old irrationality), or not to apply to expertise not claiming to be "scientific" (thus letting in ougi board practitioners), or to allow "acceptance in the particular field in which it belongs" to be defined by the practitioners of that field independent of empirical standards of verification, thus apparently letting in palmistry, dowsing, and handwriting identification expertise and leading to the ironic outcome that the best offerings of the worst fields will have a better chance of gaining admittance to a court than the merely pretty good offerings of the best fields (since the standard of the particular field is the criterion, and standards will vary across fields). The most rigorous and vigorous fields will be the slowest to grant "general acceptance" to new ideas, and the sloppiest fields will be the quickest to do so. See Saks, *Expert Testimony Before the Bench*, 90 TECH. REV. 42 (1987). In the other direction, the *Frye* test can be parodied to require such universal acceptance that one dissenter in a field will lock information out of the courtroom (a favorite strawman of people who claim to be opponents of the *Frye* test). For discussion of the current exegetical debate on how to slice the *Frye* salami, see C. MCCORMICK, *supra* note 15, § 203, at 604-10; Gianelli, *supra*, at 1228-50; McCord, *supra* note 13, at 82-88; McCormick, *supra*, at 879-916.

<sup>183</sup> Handwriting expertise has not only not been validated, it also has virtually no academic base. No accredited college or university currently offers any program leading to certification or degree. "Training" is done by apprenticeship or self-study, and the field has no uniformly recognized governing body that monitors those who identify themselves as experts. See Willard, *When and How to Use an Examiner of Disputed Documents*, 29 PRAC. LAW., Mar. 1, 1983, at 27, 28-29. The only evidence of academic consideration of this area is the Inbau study, *supra* note 26, and the recent beginnings of statistical work begun in England noted *supra* note 31.

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. First, trial courts are as a rule very conservative. They are likely to continue admitting such evidence just because it has always been admitted, at least within living memory. Further, courts are cautious bodies, and may not be sure of the implications of acting reasonably in this regard. The obvious consequences are great enough. If they knock out handwriting identification expertise, no such witness will ever be allowed to testify, and certain cases depending on the identification of disputed documents may become unprovable. Beyond this, there is the less direct but more radical fear concerning what other forms of currently accepted expertise might be equally vulnerable to attack on reexamination. For these reasons, it is incumbent upon those interested in fostering as rational a response as possible to offer some alternatives for control of handwriting identification testimony short of complete exclusion. Are there less threatening alternatives for those who seek reason, but seek it timidly? In the sections that follow we offer two alternatives, which arise at two different stages of the proceedings. The first, and preferable, alternative applies control at the point at which the expert is first exposed to the documents in the case for the purpose of comparison. The second, and perhaps more saleable, alternative subjects the expert to competency testing on cross-examination.

## VI. SUBMISSION OF DISPUTED DOCUMENTS AND EXEMPLARS UNDER A BLIND-TEST REGIME

It is reasonably clear that the asserted theoretical structure invoked by handwriting experts cannot lay claim to the name of science and that there is no reason to believe that practitioners as a group can do what they claim. Still some fraction of these practitioners may, for reasons unrelated to the theories they invoke to "explain" their conclusions,<sup>184</sup> be able to perceive and process accurately in their minds individualizing characteristics of handwriting significantly better than you or I. There are, after all, some people who are apparently better than others at seeing resemblances, identifying patterns, or discerning flavors, colors, or musical tones. Such skills do not easily lend themselves to any form of rationally persuasive "explanation" by their possessors. Such persons' conclusions might perhaps be useful, if we knew which they were. They would possess a kind of "black box" skill. We

---

<sup>184</sup> People tend to regard the assertion of "reasons" for a conclusion as a virtue. See *supra* note 174 & text preceding. In some cases, however, such as astrology or numerology, the reasons are merely a form of mystification.

(and they) wouldn't know exactly why it works, but we could still, at least theoretically, create ways to verify that it in fact works.<sup>185</sup> This is no less true for handwriting identification than for other such skills.<sup>186</sup>

We have already introduced the analogy of handwriting identification to eyewitness identification,<sup>187</sup> and we must now expand this analogy in preparation for what is to come. If there is some subgroup among the purported "experts" who can accurately identify handwriting, they would appear to be doing so in ways analogous to the average person's ability to recognize faces. The average person could describe only very imperfectly the features of a face that rendered it unique, but obviously recognition can be done much better than description. We also know, however, that the ability to recognize accurately is subject to many distorting conditions, some of which are well understood and some of which less well.<sup>188</sup> For example, research has found that eye-

---

<sup>185</sup> The notion of such testing has arisen in other judicial settings, in other cultures. In 1824, Shaka, Chief of the Zulus, reported that the royal kraal had been smeared with blood by unknown treasonous subjects who wished to put an evil spell on him. All 30,000 of his male subjects were called to assembly at the Umlhatuzana River near what is now Natal, South Africa. Shaka's Chief Witch, Nobela, and 149 assistants, then went through the crowd and fingered the culprits by magic. Shaka then revealed that he had made up the blood smearing episode as a test of the powers of his Royal Witches. Nobela and all 149 assistants were executed. See E. RITTER, *SHAKA ZULU: THE RISE OF THE ZULU EMPIRE* 217-54 (1955).

<sup>186</sup> Consider wine tasting. Let us assume a case in which a person is charged with fraud for having sold wine asserting it to be Chateau Lafitte 1966 when, according to the charge, it was not. An asserted expert in the identification of wine by taste is called by the prosecution. Should the person be allowed to testify? Some people claim such a skill, but surely their own assertion should not be sufficient to allow their testimony. It is not simply that they might be conscious humbugs. People love to believe they have wine tasting skills they do not possess. Such illusions are harmless enough when they are displayed, even convincingly displayed, in ordinary social life, but nobody ought to be subject to the determination of legal rights without greater validation than the assertion of the practitioner. To go one step further, it should not be enough even if there happens to be a society of such practitioners who all assert the skill, and affirm (collegially) each other's skill. Nor should it be enough if the society publishes a journal extolling the supposed skills of its members. Even if the proposed witness can reinforce (to herself or to her listeners) an appearance of skill by the invocation of a technical sounding vocabulary referring to an aroma of violets and blackcurrants, this is still not enough to establish validity. Either the proposed witness must be applying techniques that are replicable, and therefore available for validation to all who would take the trouble (which is not the case), or some blind testing validation would appear to be an absolute necessity.

<sup>187</sup> See *supra* note 83. The analogy is not new. See, e.g., *Handwriting as Evidence*, 60 ALB. L.J. 45, 45 (1899) (noting that identifying an individual's handwriting is similar to identifying his facial features).

<sup>188</sup> On this point, see reviews of the eyewitness accuracy literature in B.R. CLIFFORD & R. BULL, *THE PSYCHOLOGY OF PERSON IDENTIFICATION* (1978); E. LOFTUS, *EYEWITNESS TESTIMONY* (1979), A.D. YARMEY, *THE PSYCHOLOGY OF EYEWITNESS TESTIMONY* (1979); Penrod, Loftus & Winkler, *The Reliability of Eyewitness Testimony: A Psychological Perspective*, in *THE PSYCHOLOGY OF THE COURTROOM* 119 (N.L. Kerr & R.M. Bray eds. 1982).

witness accuracy varies as a function of frequency of exposure to the person to be identified, exposure time, event complexity, event stressfulness, expectations preceding exposure, manner of processing of the stored information, amount of time of information storage, characteristics of the face to be identified (of which the most commonly discussed feature is same vs. cross-racial identification), and influences during the retrieval stage.<sup>189</sup>

In order to eliminate, to the extent possible, these distorting artifacts, the law requires identification of strangers in important situations to be done in an appropriately designed "line-up." We have already touched on the characteristics of this method.<sup>190</sup> There must be enough options presented for the results to be informative, and the differences must not be suggestively obvious.<sup>191</sup> And, at least in criminal cases, this is not simply a matter of caution but of constitutional due process.<sup>192</sup>

The analogy between handwriting and eyewitness identification has its limits: The need for a proper "line-up" when questioned documents are presented to purported experts is actually even *greater* than it is when suspects are presented to eye-witnesses. We have some

---

<sup>189</sup> Because experimental psychologists tend to study more subtle phenomena, they have not looked at gross attempts at influence such as suggestions by police that they have the correct person in a lineup. Researchers have, however, found that even subtle influences can exert dramatic effects upon accuracy, among them: types of questions used to elicit witness descriptions, information contained in the question asked, and type of instructions used to elicit identifications in a lineup. "The results from these lineup studies illustrate the fact that relatively subtle differences in instructions can have rather striking effects on eyewitness performance." Penrod, Loftus & Winkler, *supra* note 188, at 141.

Somewhat analogous work by a forensic scientist found that the accuracy of conclusions from the examination of human hair varied as a function of the manner in which the samples were presented to the examiner (the traditional paired comparison of a questioned hair with a known exemplar versus a lineup style presentation). See Miller, *Procedural Bias in Forensic Science Examinations of Human Hair*, 11 L. & HUM. BEHAV. 157, 159-62 (1987). This study is a beginning and a rarity. In general, the forensic science profession has not undertaken such research to determine the optimal conditions in which to do their work and maximize their accuracy. One can only be astounded at the volume of research on eyewitness accuracy and the paucity of parallel work on forensic science accuracy.

<sup>190</sup> See *supra* note 78.

<sup>191</sup> Again, research psychologists have done a considerable amount of work on what really does or does not constitute a fair test of the witness. This research has developed an empirical technique for assessing the "functional size" of a lineup as distinct from its "nominal" size (that is, the size of a lineup adjusted for the fact that some people in it can be ruled out even by someone who did not witness the crime), and the probability of an erroneous identification can be calculated from the lineup's functional size. See Wells, Leippe & Ostrom, *Guidelines for Empirically Assessing the Fairness of a Lineup*, 3 L. & HUM. BEHAV. 285, 288-90 (1979).

<sup>192</sup> Flowing of course from *United States v. Wade*, 388 U.S. 218 (1967), and its progeny. *Wade* is as much an apogee for the review of the rationality of evidentiary practices as *Barefoot v. Estelle*, 463 U.S. 880 (1983), is a nadir. See *infra* note 212.

"clinic of life" reason to believe that the average eyewitness can make some kinds of identifications, but we have no real reason to believe any handwriting expert can make *any* identifications, at least not any laypersons could not make from the available data. At best, there is a hypothetically possible (and apparently small)<sup>193</sup> subset of such "experts" who may possess the skill mixed in with what is probably a majority who don't. Beyond that, there is no reason to believe that handwriting "experts" are not just as subject to problems of suggestiveness as everyone else. Whatever else is said about their skill, it is not something like a chemical titration that is largely independent of influence by suggestion. There is, in fact, some anecdotal<sup>194</sup> reason to believe that handwriting "experts" are very subject to suggestion,<sup>195</sup> which makes the non-blind normal method of presenting them with a problem close to a not-very-amusing joke. Thus, those courts that refuse to exclude such testimony should be more receptive to requiring that all handwriting identifications be the product of appropriately designed and presented blind tests. In these tests the exemplars of the suspected candidate for authorship ought to be presented to the expert only in the coded company of a meaningful number of other appropriately similar exemplars from people who clearly did not write the contested document.

Two of the authors have been involved in three cases in which this approach was undertaken, with an interesting set of results, two of which are directly relevant here. In the first case, the prosecutor thought the proposal sounded fair and agreed to it. He then discovered that no handwriting expert he could find would agree to undertake such a test.<sup>196</sup> In the second case the court decided that, under binding decisions of the New Jersey Supreme Court, it was obligated to order the exemplars produced and had no discretion to order them only upon blind test structuring, but it told the prosecutor that if blind test struc-

---

<sup>193</sup> See *supra* notes 14-80 and accompanying text.

<sup>194</sup> We realize that we are here stooping to conquer anecdotal claims of expertise with anecdotal criticism. Given the paucity of the empirical record, a paucity to which handwriting analysts' apparent reluctance to be tested has contributed, we have no alternative.

<sup>195</sup> According to reports of FBI Agent Thomas Sisk and others, Albert S. Osborn had tentatively found that, although there were similarities, Richard Hauptmann's handwriting did not match the Lindberg baby ransom notes. Less than an hour after finding out that a large part of the ransom money had been discovered in Hauptmann's garage, he called with a revised opinion. See L. KENNEDY, *THE AIRMAN AND THE CARPENTER* 178-83 (1985). For an earlier case illustrating the problems of suggestion, see *R. v. Silverlock*, 2 Q.B. 766, 767 (1894).

<sup>196</sup> It is traditional to refer to the docket number of the case here. Since no charges were ever filed, however, there is none; and as it would be wrong to give the client's name under the circumstances, the reader will have to take our word for it.

turing was not undertaken the court might ultimately rule inadmissible any opinion rendered in a non-blind presentation.<sup>197</sup> It appears that demanding a blind testing regime from the start at least gives courts pause, and may make experts less willing to be involved in the case on a business as usual basis. Ultimately, the courts ought and may be willing to adopt this approach if they are unwilling to embrace total exclusion.

## VII. TESTS OF SKILL DURING CROSS EXAMINATION

Suppose that a significant number of courts refuse to exclude handwriting identification "expert" testimony totally, and further believe themselves without authority either to order exemplars only in the context of a properly designed blind test or to exclude the resulting non-blind identification. What response can be made? Here the third case we have been involved in<sup>198</sup> becomes relevant. In that case, we made our usual motion resisting the production of exemplars except in the context of a properly designed blind test. The court ruled that it had no authority to require such a testing regime as a condition of the production of the exemplars but further ruled that the expert could be subjected to such competency tests on cross examination.<sup>199</sup> The expert rendered that issue moot by then finding the results of his non-blind examination of the disputed documents and exemplars to be "inconclusive." The prospect of undergoing properly designed blind tests of asserted skill seems once again to have made the expert reluctant to participate in the usual way.<sup>200</sup>

It is clear that the opponent of the witness has a right to impeach by providing information to the jury, through cross examination or otherwise, that would rationally affect how much to discount the face value of the witness' testimony. This might, of course, include cross examining the witness on the information revealed in this article. But more important, it ought to include subjecting the witness to various

---

<sup>197</sup> Unreported oral opinion of Judge D'Annunzio in *State v. Padgett*, Indictment No. 308-M79 (N.J. Superior Court, Hunterdon County Dec. 1980) (observed by Professor Risinger).

<sup>198</sup> Unreported oral opinion in *People v. Barrows*, No. 8280/84 (N.Y. Sup. Ct., N.Y. County May, 1985) (the "Mayflower Madam" case).

<sup>199</sup> *See id.* This oral opinion was unpublished. Professor Denbeaux was present in the courtroom.

<sup>200</sup> There is reported evidence supporting this view. In *Heller v. Murray*, 112 Misc. 2d 745, 442 N.Y.S.2d 348 (N.Y. Civ. Ct. 1981), *aff'd*, 118 Misc. 2d 508, 464 N.Y.S.2d 391 (N.Y. App. Term. 1983), the handwriting witness refused to take a test during cross-examination that the judge had ruled appropriate. The court ultimately ruled that the signature in question was not that of the person who was identified as the author by the "expert."

fairly designed tests of the witness' abilities, to be performed during cross examination. The right to do so would seem to be protected by both due process interests and, in a criminal case, the confrontation clause.

We not only believe such a right exists, we believe it is supported by virtually all the published judicial authority in the past fifty years and by the main gurus of handwriting identification, such as Osborn, Wigmore, and Ordway Hilton.<sup>201</sup> It also appears, however, that in practice such testing is rarely undertaken. The reason for this rarity seems to be another unexamined legacy of the 19th century common law.

To expand on the preceding paragraph, the reader will remember that during the last thirty-five years of the 19th century, as courts were coming to admit such "expertise" even though they manifested great hostility and contempt for its dependability, a number of courts refused to depart from their reading of the common law rule forbidding "extrinsic" exemplars.<sup>202</sup> Some forbade them entirely and some allowed them only to the minimum amount necessary to lay the foundation for the "expert's" testimony.<sup>203</sup> The effect was to render fair tests of skill unworkable in the courtroom, because such tests would almost inevitably involve carefully selected and prepared documents that were "extrinsic."<sup>204</sup> Fairly early on, this effect provoked a reaction against applying the common law exclusionary rule to limit cross-examination of experts. Beginning in Michigan in 1888,<sup>205</sup> spreading to Georgia (1890),<sup>206</sup> Iowa (1894),<sup>207</sup> and Connecticut (1902),<sup>208</sup> and culminating in the previously quoted New York Court of Appeals case of *Hoag v. Wright*,<sup>209</sup> some courts vigorously championed the right of the opponent of such a putative handwriting expert to subject him to tests on cross examination using otherwise "extraneous" exemplars. More important, in the last half century, though the reported decisions have been few,

---

<sup>201</sup> See A. OSBORN, *supra* note 13, at 36-38; 7 J. WIGMORE, *supra* note 15, § 2015; Hilton, *Cross Examination of a Handwriting Expert by Test Problem*, 13 RUTGERS L. REV. 306 (1958).

<sup>202</sup> See *supra* note 98 and accompanying text.

<sup>203</sup> See *id.*

<sup>204</sup> The roots of the exclusion of such tests actually predated "experts." The first cases forbade subjecting recognition witnesses to such tests. In doing this the courts may have been influenced by the intuition that recognition witnesses would fail such tests and thus the whole edifice of conventional authentication would crumble. See *Griffits v. Ivery*, 11 A. & E. 322 (Q.B. 1840).

<sup>205</sup> *Johnston Harvester Co. v. Miller*, 72 Mich. 265, 40 N.W. 429 (1888).

<sup>206</sup> *Traveler's Ins. Co. v. Sheppard*, 85 Ga. 751, 12 S.E. 18 (1890).

<sup>207</sup> *Browning v. Gosnell*, 91 Iowa 448, 59 N.W. 340 (1894).

<sup>208</sup> *Brown v. Woodward*, 75 Conn. 254, 53 A. 112 (1902).

<sup>209</sup> 174 N.Y. 36, 66 N.E. 579 (1903).



more courts have accepted such testing than have rejected it;<sup>210</sup> and the last reported decision rejecting it was in 1940.<sup>211</sup> Though the most recent precedent in some jurisdictions may be a restrictive nineteenth or early twentieth century decision,<sup>212</sup> it is doubtful that that decision would, or even could, be followed today if the issue were explicitly faced and reconsidered. Rather, it is more accurate to say that the modern position, when courts actually think their way through to one, is that such testing is allowable. It is difficult to explain with any certainty why, in the face of this position, even the attempted use of such tests is so uncommon.<sup>213</sup> At the very least any such "expert" allowed to testify ought to be subject to some appropriately designed test on cross examination.

### CONCLUSION

We have reached the end of our main topic and it only remains for us to ask if there are any greater lessons to be learned from the law's treatment of handwriting expertise. One is tempted to say that as a general proposition, the law does not yet know how to deal with science,<sup>214</sup> or with things asserting themselves to be science.<sup>215</sup> On the one

---

<sup>210</sup> The leading case is *State v. Bulna*, 46 N.J. Super. 313, 134 A.2d 738 (App. Div. 1957), *aff'd*, 27 N.J. 93, 141 A.2d 529 (1958). *See also* *Heller v. Murray*, 112 Misc. 2d 745, 447 N.Y.S.2d. 348 (N.Y. Civ. Ct., 1981), *aff'd*, 118 Misc. 2d 508, 464 N.Y.S. 2d 391 (N.Y. App. 1983).

<sup>211</sup> *State v. Maxwell*, 151 Kan. 951, 102 P.2d. 109 (1940) (relying on a previous 19th century Kansas precedent).

<sup>212</sup> *See* 7 J. WIGMORE, *supra* note 15, § 2015 n.2.

<sup>213</sup> We have a plausible theory. First, it seems likely that Wigmore and Osborn sold most lawyers on the basic reality of the expertise during the first third of this century. If one assumes this belief, then an in-court test is risky, since passing it may be devastating. Second, designing a proper test is difficult and requires outside sources of exemplars. In the press of business and in the face of perceived risks, most lawyers probably come to court without preparing a test to give. Third, it is easier and safer to counter the other side's expert with one of your own than it is to prepare and administer an in-court test. If there is any reality to the last speculation, it represents an unstated conspiracy of adversaries to continue with irrational evidence irrationally countered, rather than pressing for the validly meaningful. Put that in your "search for truth" model.

<sup>214</sup> *Ius not calculat*. One may speculate that many people go to law school partly because it requires little math.

<sup>215</sup> A recent article in the nation's leading scientific journal, *see* Black, *Evolving Legal Standards for the Admissibility of Scientific Evidence*, 239 SCIENCE 1508 (1988), claims to discern a "trend toward more thorough judicial review of scientific claims." What we perceive, however, is paradigm chaos. Prior to the late 1960s there was a general acceptance of whatever "the *Frye* test" was taken to mean, even though a lot of unvalidated testimony was allowed in practice. Since the late 1960s there has been a series of decisions hostile to or myopic concerning requirements of empirical validation, just as there has been a series of opinions sensitive to and knowledgeable about what standards of empirical validation are and why they are important. The

hand there is the *Frye* test and its variants, which seem to incorporate a requirement of validity testing from the scientific community as a condition of admissibility of evidence claiming the mantle of science. On the other hand, the *Frye* test has been criticized as denying the law the benefits of new developments and has been grudgingly construed to apply only to forms of expertise that are both centrally scientific and novel.<sup>216</sup> Thus, alleged areas of expertise, whose practitioners eschew the label "scientific", are sometimes accepted without evidence of validity; so are forms of expertise that predated the *Frye* test. A lot of unvalidated nonsense is allowed into court under circumstances that look suspiciously close to the old gentlemen's agreement from the days of Langdellian-Wigmorean "juridical science": "I won't look too closely at your claims to 'scientific' expertise if you don't look too closely at mine."

---

most famous instances within the "hostile" or "myopic" series are *Coppolino v. State*, 223 So. 2d. 68 (Fla. Dist. Ct. App. 1968), *cert. denied*, 399 U.S. 927 (1969); *U.S. v. Stifel*, 433 F.2d. 431 (6th Cir. 1970), *cert. denied*, 401 U.S. 994 (1971); and *Barefoot v. Estelle*, 463 U.S. 880 (1983). They have been generally disastrous.

In *Coppolino*, testimony by Dr. Milton Halpern concerning a newly developed and unvalidated test for the metabolic byproducts of injected succinylcholine chloride (a theretofore untraceable poison) was allowed in, despite Halpern's admission that the data supporting the test were not in publishable condition. *Coppolino* was convicted. The results were never published and the process has never been replicated. (This lack of validation may or may not have had something to do with *Coppolino's* later parole. See *NEWSWEEK*, May 7, 1979, at 16.)

In *Stifel*, testimony was allowed in concerning unvalidated applications of neutron activation analysis. *Stifel* was convicted. The conviction was affirmed by the Sixth Circuit in an opinion much celebrated by the proponents of loose admissibility for expert testimony associated, perhaps unfairly, with the late Charles T. McCormick. See generally Gianelli, *supra* note 182, at 1232-1235 (discussing the confusion generated by McCormick's view on admissibility). Later *Stifel's* conviction was vacated based on data that showed that the test could not do what the witness testified it had done. See *U.S. v. Stifel*, 594 F. Supp. 1525 (N.D. Ohio 1984). *Stifel* later went to Cleveland State Law School and in 1987 won one of the national moot court competitions, the Jerome Prince Evidence Moot Court Competition at Brooklyn Law School. He did such a good job that Judge Weinstein, one of the judges in the final who was apparently not cognizant of his background, referred to him in his oral evaluation as "a natural born prosecutor", see Weinstein, Mansfield, Abrams & Berger, *EVIDENCE: CASES AND MATERIALS* 386 (8th ed. 1988).

Finally, the Supreme Court in *Barefoot* held that psychiatric testimony concerning dangerousness could be admitted in the sentencing hearing of a death penalty case without violating due process requirements, even though there was overwhelming empirical evidence that such dangerousness predictions are *invalid* and at least twice as likely to be wrong as right. As the *Science* article referred to above concedes, *Barefoot* is a low point in the struggle to hold the trial system to its own stated ideals concerning "search for truth" See Black, *supra*, at 1509 ("*Barefoot* . . . is perhaps the most egregious example of how courts refuse to accept the limitations of psychiatry, and refuse to hold psychiatrists to the accepted standards of their own profession."). Indeed, we have yet to find a single word of praise for, or in defense of, *Barefoot* in the literature of either science or law.

<sup>216</sup> See *supra* note 179.

Finally, beyond the dark corners of inherited expertise, we may perhaps draw a broader lesson still. The "search for truth" model<sup>217</sup> of our procedural system, which comprises much of its official ideology, cannot account for many phenomena in the real world of litigation. Many have therefore proposed an alternate lens for viewing the reality of litigation, the "judicial combat" model.<sup>218</sup> The proponents of this view assert that what we actually do (as opposed to what we say we do) is seek, not a result that mirrors the law applied to the facts as they were, but a result that is the product of a fair and satisfying contest, whatever the facts may have been.<sup>219</sup> If this latter view is not suffi-

---

<sup>217</sup> See FED. R. EVID. 102 (1988) "These rules shall be construed to secure fairness in administration, elimination of unjustifiable expense and delay, and promotion of growth and development of the law of evidence to the end that the truth may be ascertained and proceedings justly determined." See also *Williams v. Florida*, 399 U.S. 78, 81-82 (1970) (The adversary system is designed "to enhance the search for truth in the criminal trial by ensuring both the defendant and the state ample opportunity to investigate certain facts crucial to the determination of guilt or innocence."); *Riley v. Goodman*, 315 F.2d 232, 234 (3d Cir. 1963) (citation omitted) ("We have long abandoned the adversary system of litigation which regards opposing lawyers as players and the judge as a mere umpire whose only duty is to determine whether infractions of the rules of the game have been committed. A trial is not a contest but a search for the truth so that justice may properly be administered.").

<sup>218</sup> See, e.g., 7 J. WIGMORE, *supra* note 15, § 1845.

[T]he common law, originating in a community of sports and games, was permeated essentially by the *instincts of sportsmanship*. This has had both its higher and lower aspects. On one hand, it has contributed a sense of fairness, of chivalrous behavior to a worthy adversary, of carrying out a contest on equal and honorable terms. . . . On the other hand, it has contributed to lower the system of administering justice, and in particular of ascertaining truth in litigation, to the level of a mere game of skill or chance.

*Id.* Some people can't seem to separate the two visions. Consider, for example, Bandes, *Taking Some Rights Too Seriously: The State's Right to a Fair Trial*, 60 S. CAL. L. REV. 1019 (1987).

The classic view of the trial is that it is a "search for truth." In this view, the adversary system is a contest between equals. The role of the prosecutor is to obtain a conviction, while the role of the defense attorney is to obtain an acquittal. If each carries out his role, then the truth will emerge. The goal of this model is to arrive at an accurate result. The search for truth model demands that the "scales must be evenly held," so that the parties may be equally armed for adversarial combat.

*Id.* at 1037 (footnotes omitted);

<sup>219</sup> This outlook does not of course exhaust the alternate lenses that have been proposed. See, e.g., Graham, *supra* note 147, [proposing (perhaps partially tongue in cheek—it is difficult to tell, as the piece mixes styles from the critical legal studies school with the "gonzo" style of journalist Hunter Thompson) a model of Legal Irrationalist Disbelief, in which a trial is a political exercise and the ideas of fair fight and of rationality are only tools to be used when the other side can be conned into letting them help you.]. Even in a system that is internally consistent and coherent (which ours by no means is), the existence of mutually inconsistent lenses by which to view and account for what is happening is probably inevitable. See Leff, *Law and*, 87 YALE L.J.

ciently cynical as an account of many of the details of actual practice, we are afraid that the handwriting expertise phenomenon, and others,<sup>220</sup> lead us to propose still another model: the “exorcism of ignorance” model. This model holds that, under the stress of having to wrestle with important types of facts about which there is no good evidence, the system will invite the creation of a proxy for rational knowledge, a form with the appearance of evidence but no rational content, to be used in a ritual exorcism of an ignorance we cannot bear. If this in fact accounts for some of what happens in the real world, we can only hope that the legal system’s conscious commitment to the “search for truth” model will lead it to respond honorably when the imperfection in our knowledge is finally pointed out in ways that cannot be avoided.

---

989 (1978).

<sup>220</sup> The recent struggles of the litigation system to deal with the problems of proof in child molestation and abuse cases immediately spring to mind. Charles Nesson has recently posited, in very tentative form, a similar mechanism in the procedural law whereby the “search for truth” is sometimes compromised to enhance the acceptability of results. See Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357 (1985).

## APPENDIX 1—LITERATURE SEARCHED

In conducting the literature search, we started with the books listed by Maureen Casey Owens, Chief Document Examiner for the Chicago Police Department, in response to our inquiry. *See supra* note 27. We added to that list the relevant books in English from Osborn's critical bibliography appended to his 1929 second edition of *Questioned Documents*. We went on to examine other books published since 1929, including the relevant chapters in general works on forensic science. The books examined were:

- D. AMES, *AMES ON FORGERY* (1899).
- D. BLACKBURN & W. CADDELL, *THE DETECTION OF FORGERY* (1909).
- L. CAPUTO, *QUESTIONED DOCUMENT CASE STUDIES* (1982).
- C. CHABOT, *THE HANDWRITING OF JUNIUS, PROFESSIONALLY INVESTIGATED* (1871).
- J. CONWAY, *EVIDENTIAL DOCUMENTS* (1959).
- LAW, MEDICINE, AND FORENSIC SCIENCE (W. Curran & E. Shapiro eds. 3d ed. 1982).
- P. FRAZER, *A MANUAL OF THE STUDY OF DOCUMENTS* (1894).
- P. FRAZER, *BIBLIOTICS, OR THE STUDY OF DOCUMENTS* (3d ed. 1901).
- P. GIANELLI & E. IMWINKELRIED, *SCIENTIFIC EVIDENCE* (1986).
- W. HAGEN, *DISPUTED HANDWRITING* (1894).
- W. HARRISON, *SUSPECT DOCUMENTS: THEIR SCIENTIFIC EXAMINATION* (1958).
- SCIENTIFIC AND EXPERT EVIDENCE* (E. Imwinkelried ed. 1981).
- C. LEE & R. ABBEY, *CLASSIFICATION AND IDENTIFICATION OF HANDWRITING*.
- C. MITCHELL, *DOCUMENTS AND THEIR SCIENTIFIC EXAMINATION* (1922).
- A. MOENSSSENS, R. MOSES, & F. INBAU, *SCIENTIFIC EVIDENCE IN CRIMINAL CASES* (1973).
- A.S. OSBORN, *THE PROBLEM OF PROOF* (1922).
- A.S. OSBORN, *QUESTIONED DOCUMENTS* (2d ed. 1929).
- A.S. OSBORNE & A.D. OSBORNE, *QUESTIONED DOCUMENT PROBLEMS* (2d ed. 1946).
- A. PETERSON, E. FABRICANT & K. FIELD, *CRIME LABORATORY PROFICIENCY TESTING RESEARCH PROGRAM: FINAL REPORT* (1978) (U. S. Government Printing Office) (Sup't Docs. mo. cat. no 79-6238).
- A. QUIRKE, *FORGED, ANONYMOUS, AND SUSPECT DOCUMENTS* (1930).

FORENSIC SCIENCES (C.H. Wecht ed. 1986).

The search of the periodical literature was somewhat more difficult. Ms. Casey Owens identified the following journals as the main ones in which articles on handwriting identification appear:

The Journal of Forensic Sciences

Forensic Science International

The Journal of Police Science and Administration

The Journal of the Forensic Science Society

The Canadian Society of Forensic Science Journal

In addition to these, it appears to us that the bulk of published material on handwriting identification has appeared in *The Journal of Criminal Law, Criminology and Police Science*, and its successor, *The Journal of Criminal Law and Criminology*. This journal also contains an extensive section of abstracts of articles from other forensic science journals worldwide, such as the *Australian Journal of Forensic Science*.

Because most forensic sciences have little academic base, these materials were hard to obtain. The initial gathering was done by Dr. Saks while he was on the faculty of Boston College. Even using the Boston (metropolitan) library consortium, The Harvard Law Library, and the Library of the Director of the Massachusetts State Police Laboratory, some were not obtainable in the Boston area. Nevertheless, we have persevered, and have now reviewed all or substantially all of the issues of these journals, including every article cross-referenced in any other article which appeared to be a potential source of empirical data. The results are reflected in the text.

The vast majority of the handwriting analysis "studies" in these publications consist of anecdotal observations, hypothesis creation, and speculation.

## APPENDIX 2

CORRELATION OF ACCURACY WITH EXPERIENCE,  
TIME EXPENDED, AND CERTIFICATION: FSF TESTS  
FOR 1984 THROUGH 1987

## 1984

Only time-on-task data were collected this year. Those who got the analysis right spent a mean of 3.219 hours on the task and those who got it wrong spent a mean of 2.375 hours — an apparent difference which, with this sample size and variation in time spent, was not statistically significant ( $t=0.668$ ;  $df=20$ ;  $p=0.512$ ). (Meaning that chance cannot safely be excluded as a cause of the apparent differences. By conventional tests of significance, a chance difference can be ruled out only if the probability of error in so concluding is less than 5%, that is,  $p<0.05$ .)

For the raw data, *see* 1984 FSF STUDY, *supra* note 34, at 6.

## 1985

Again, only time-on-task data were collected. Comparing those who got correct answers (mean time 4.57 hours) to those who got wrong answers (mean time 2.31 hours) to those who could not reach a conclusion (mean time 2.12 hours), the apparent differences again were not statistically significant ( $F=1.77$ ;  $df=2,28$ ;  $p=0.188$ ). Comparing only those getting right answers and those getting wrong answers produces the same non-significant result ( $F=1.70$ ;  $df=1,21$ ;  $p=0.206$ ).

For the raw data, *see* 1985 FSF STUDY, *supra* note 35, at 10.

## 1986

This year data on years of experience in document examination and percentage of caseload which consists of document examination work were collected along with time-on-task. In regard to time-on-task, comparing producers of right answers (mean time 2.39 hours), versus wrong answers (mean time 2.80 hours) versus inconclusive answers (mean time 3.22 hours) again produced nonsignificant differences ( $F=0.362$ ;  $df=2,27$ ;  $p=0.699$ ). So did a comparison of only right versus wrong groups ( $F=0.208$ ;  $df=1,19$ ;  $p=0.654$ ).

In regard to years of experience, a marginally significant result emerged. Those who got the answer right averaged 18.33 years of experience compared with 10.43 for those who were wrong and 11.07 for those who were inconclusive ( $F=2.59$ ;  $df=2,25$ ;  $p=0.095$ ). A direct

comparison of those right versus those wrong yielded a similarly marginally significant difference ( $F=4.11$ ;  $df=1,19$ ;  $p=0.057$ ). These barely significant differences should not be permitted to conceal the variation that still exists. Examiners with as much as twenty years of experience still managed to produce wrong answers, and some with as little as five years experience managed to produce correct answers.

Finally, when those whose caseloads are exclusively document examination were compared with those whose caseloads were mixed with other kinds of cases, no difference was found. (Comparing right to wrong to inconclusive,  $\chi^2=0.860$ ;  $df=2$ ;  $p=0.651$ . Comparing right to wrong, using Fisher's Exact Test,  $p=0.590$ .)

For the raw data, *see* 1986 FSF STUDY, *supra* note 36, at 11-12.

### 1987

This year's test included all the variables of the previous year, plus an indication of whether the examiner was certified by the American Board of Forensic Document Examiners. Because the test this year was so easy that only one clear error was produced, our statistical comparisons had to be limited to those right versus those inconclusive.

Time-on-task revealed that those who produced right answers took less time (a mean of 2.75 hours) than those who could not reach a conclusion (a mean of 5.00 hours) ( $t=2.50$ ;  $df=26$ ;  $p=0.019$ ). This difference may mean only that those who couldn't figure out the answer did not give up without trying.

Years of experience again produced a marginally significant difference, with those who got it right averaging 13.36 years in the business compared with inconclusive respondents' averaging 7.05 years ( $t=1.96$ ;  $df=22$ ;  $p=0.063$ ). This result suggests that those who work at handwriting analysis for enough years do get slightly better at it. On the other hand, the causation may run in a different direction: for example, those with greater confidence in their analyses may tend to stay in the profession longer. In any case, generalization may mislead more than it helps since the variation in both groups is so large. For example, at one lab four examiners, including a supervisor, *averaged* 17.5 years' experience among them and could not come up with the right answer, while other examiners with as little as two years' experience were able to do so.

Having a caseload composed exclusively of document examination again made no difference. (Those with exclusive caseloads were right 53% of the time, while those with mixed caseloads were right 54% of the time; by Fisher's Exact Test,  $p=0.724$ .) And those certified by the American Board of Forensic Document Examiners were insignificantly



more likely to get the answer (60%) than those without certification (50%) (by Fisher's Exact Test,  $p=0.456$ ).

For the raw data, *see* 1987 FSF STUDY, *supra* note 37, at 4-5.

## APPENDIX 3

DATE OF FORMAL ACCEPTANCE OF HANDWRITING  
EXPERT TESTIMONY BY EACH STATE  
(IN CHRONOLOGICAL ORDER)

<u>State</u>	<u>Date</u> <sup>1</sup>	<u>Instrument of Acceptance</u>
Louisiana	1808	By statute. <i>See Sauve v. Dawson</i> , 2 Mart. 202 (1812) (citing LA. CIV. CODE ANN. art. 226 (1808), <i>reprinted as</i> LA. CIV. CODE ANN. art. 2245 (West 1952). <i>See also</i> LA. CIV. CODE ANN. art. 1838 (West 1987) (current statute).
Massachusetts	1836	By decision. <i>See Moody v. Rowell</i> , 34 Mass. (17 Pick.) 490 (1836).
Ohio	1850	By decision. <i>See Hicks v. Person</i> , 19 Ohio 426 (1850).
Iowa	1851	By statute. <i>See</i> IOWA CODE § 2404 (1850-1851). <i>See also</i> IOWA CODE ANN. § 622.25 (West 1950) (current statute).
Maine	1851	By decision. <i>See Sweetser v. Lowell</i> , 33 Me. 446 (1851).
New Hampshire	1852	By decision. <i>See Bowman v. Sanborn</i> , 25 N.H. 87 (1852).
Florida	1861	By statute. <i>See</i> DIG. FLA. LAWS ch. 101, § 16 (McClellan 1822-81). <i>See also</i> FLA. STAT. ANN. § 92.38 (West 1979) (current statute).
Oregon	1862	By statute. <i>See</i> OR. GEN. LAWS (Civ. Code ch. 8, § 755) (1843-1872). <i>See also</i> OR. REG. STAT. § 40.505 (1953).
New Jersey	1865	By statute. <i>See Mutual Benefit Life Ins. Co. v. Brown</i> , 30 N.J. Eq. 193 (1878) (citing N.J. REV. STAT. § 19, at 381 (1877), which in turn cites "P.L. 1865, p. 832, § 1").

---

<sup>1</sup> All statute dates were taken from cases referenced, or from notes to 7 WIGMORE §§ 2008, 2016, unless otherwise noted.

Texas	1866	By statute. <i>See</i> PASCHAL's DIG. art. 3132 (1866) (cited in <i>Lawson, supra</i> note 136, at 59). <i>See also</i> TEX. R. CIV. EVID. ANN. r. 901(b)(3) (Vernon 1988); TEX. R. CRIM. EVID. ANN. r. 901(b)(3) (Vernon 1988) (current statute).
Vermont	1867	By decision. <i>See State v. Ward</i> , 39 Vt. 225, 237 (1867).
Nebraska	1867	By statute. <i>See</i> NEB. REV. STAT. § 344 (1867) (cited in <i>Grand Island Banking Co. v. Shoemaker</i> , 31 Neb. 124, 127, 47 N.W. 696, 697 (1891)). <i>See also</i> NEB. REV. STAT. § 25-1220 (1943) (current statute).
Connecticut	1869	By decision. <i>See Tyler v. Todd</i> , 36 Conn. 218, 221-22 (1869).
Indiana	1870	By decision. <i>See Chance v. Indianapolic &amp; Westfield Gravel Rd. Co.</i> , 32 Ind. 472, 474 (1870).
Missouri	1870	By decision. <i>See State v. Scott</i> , 45 Mo. 302, 305-06 (1870).
California	1872	By statute. <i>See</i> CAL. CODE CIV. PROC. § 1870(9) (1872), <i>codified at</i> CAL. CIV. PROC. CODE § 1870(9) (West 1955), <i>superseded by</i> CAL. EVID. CODE §§ 720, 800, 801, 1416 (West 1983).
Kansas	1872	By decision. <i>See Macomber v. Scott</i> , 10 Kan. 335, 340-41 (1872).
Mississippi	1874	By decision. <i>See Wilson v. Beauchamp</i> , 50 Miss. 24, 32 (1874).
West Virginia	1874	By decision. <i>See Clay v. Robinson</i> , 7 W. Va. 348, 361-63 (1874).
North Carolina	1877	By decision. <i>See Yates v. Yates</i> , 76 N.C. 142, 149-50 (1877).
New York	1878	By decision. <i>See Miles v. Loomis</i> , 75 N.Y. 288, 297-98 (1878).
Montana	1878	By decision. <i>See David v. Fredericks</i> , 3 Mont. 262, 265-66 (1878).
Delaware	1879	By statute. <i>See</i> 16 Del. Laws ch. 536 (1879), <i>codified at</i> DEL. REV. STAT. § 4230 (1915). <i>See also</i> DEL. CODE ANN. tit. 10, § 4310 (1974) (current statute).

- |                |      |   |
|----------------|------|---|
| Wisconsin      | 1881 | By statute. <i>See</i> 1881 Wis. Laws ch. 226 ch 226. <i>See also</i> DEL. CODE ANN. tit. 10 § 4310 (1974) (current statute).   |
| Virginia       | 1884 | By decision. <i>See Hanroit v. Sherwood</i> , 82 Va. 1, 16-17 (1884).   |
| Kentucky       | 1886 | By statute. <i>See</i> CIV. CODE PRAC. § 604 (1886). (cited in <i>Storey v. First Nat'l Bank</i> , 72 S.W. 318, 320 (Ky. App. 1903)). <i>See also</i> KY. REV. STAT. ANN. § 422.120 (1972) (current statute).   |
| Georgia        | 1886 | By decision. <i>See Smith v. State</i> , 77 Ga. 705 (1886).   |
| Michigan       | 1886 | By decision. <i>See Vinton v. Peck</i> , 14 Mich. 286, 294 (1886).  |
| Minnesota      | 1886 | By decision. <i>See Morrison v. Porter</i> , 35 Minn. 425, 426, 29 N.W. 54, 54-55 (1886).   |
| Utah           | 1887 | By decision. <i>See Durnell v. Sowden</i> , 5 Utah 216, 14 P. 334, 335-36 (1887).   |
| Maryland       | 1888 | By statute. <i>See</i> 1888 Md. Laws ch. 545, <i>codified at</i> MD. ANN. CODE art. 35 § b. <i>See also</i> MD. CTS. & JUD. PROC. CODE ANN., § 10-906 (1984) (current statute).   |
| Tennessee      | 1889 | By statute. <i>See</i> 1889 Tenn. Pub. Acts, ch. 22. <i>See also</i> TENN. CODE ANN. § 24-7-108 (1980) (current statute).   |
| South Carolina | 1890 | By decision. <i>See State v. Ezekial</i> , 33 S.C. 115, 116, 11 S.E. 635, 636 (1890) (dictum); <i>State v. Lyle</i> , 125 S.C. 406, 428, 118 S.E. 803, 811 (1923) (adopting <i>Ezekial</i> ).   |
| Illinois       | 1892 | By decision. <i>See Rogers v. Tyley</i> , 144 Ill. 652, 665, 32 N.E. 393, 397 (1892).   |
| Colorado       | 1893 | By statute. <i>See</i> 1893 Colo. Sess. Laws ch. 88. <i>See also</i> COLO. REV. STAT. § 13-25-104 (1987) (current statute).   |
| Pennsylvania   | 1895 | By statute. <i>See</i> 1895 Pa. Laws no.49 (P.L. 69) (cited in <i>Groff v. Groff</i> , 209 Pa. 603, 612, 59 A. 65, 68 (1904)). <i>See also</i> PA. STAT. ANN. tit. 28, §§ 161-63 (Purdon 1958), <i>repealed by</i> 1978 Pa. Laws no.53 § 2(a) (P.L. 202). |

Washington	1896	By decision. <i>See Moore v. Palmer</i> , 14 Wash. 134, 138, 44 P. 142, 144 (1896).
Idaho	1900	By decision. <i>See Bane v. Gwinn</i> , 7 Idaho 439, 447, 63 P. 634, 636 (1900).
Oklahoma	1900	By decision. <i>See Archer v. United States</i> , 9 Okla. 569, 675, 60 P. 268, 270 (1900).
Arizona	1901	By territorial statute. <i>See</i> 1901 ARIZ. REV. STAT. tit. 28, ch. 4 § 47 (¶ 2544) (1901). <i>See also</i> ARIZ. REV. STAT. ANN. Rules of Civ. Proc., 44(m) (1956) (current statute).
Rhode Island	1905	By statute. <i>See</i> Court Practice Act of 1905, § 399 (cited in <i>Municipal Ct. v. Kirby</i> , 28 R.I. 287, 288, 67 A. 8, 9 (1907), <i>reprinted</i> in 1909 R.I. Acts & Resolves ch. 292, § 47. <i>See also</i> R.I. GEN. LAWS § 19-19-17 (1985) (current statute).
South Dakota	1906	By decision. <i>See In re McClellan's Estate</i> , 20 S.D. 498, 509-10, 107 N.W. 681, 685-86 (1906).
Alabama	1907	By statute. <i>See</i> ALA. CIV. CODE ch. 84, art. 2, § 4005 (1907), though there was a conflict in previous decision. <i>Compare Moon's Adm'r. v. Crowder</i> , 72 Ala. 79 (1882) (allowing expert testimony based upon comparisons of genuine documents), <i>with Gibson v. Trowbridge Furniture Co.</i> , 96 Ala. 357, 361-62, 11 So. 365, 366 (1891) (excluding expert comparisons of handwriting entirely). <i>See also</i> ALA. CODE § 12-21-62 (1986) (current statute).
U.S.	1913	By statute. <i>See</i> 37 Stat. 683, ch. 79 (1913), <i>codified at</i> 28 U.S.C. § 1731 (1983). Previous cases were split. <sup>2</sup>
New Mexico	1917	By statute. <i>See</i> 1917 N.M. LAWS, ch. 64. <i>See also</i> N.M. STAT. ANN. § 11-901(B)(3) (1978) (current statute).

---

<sup>2</sup> *Compare Winthrop v. United States*, 127 F. 530, 535 (8th Cir. 1903) (comparison allowable under some conditions); *National Accident Soc'y v. Spire*, 78 F. 775, 777 (6th Cir. 1897) (expert testimony probably not admissible); *Smyth v. New Orleans Canal & Banking Co.*, 93 F. 899 (5th Cir. 1899) (exemplars not otherwise related to case allowed); *United States v. Mathias*, 36 F. 893, 894 (C.C.D.S.C. 1888) (expert testimony allowed).

Hawaii	1925	By territorial statute. <i>See</i> HAW. REV. L. § 2629 (1925) <i>See also</i> HAW. REV. STAT. § 626-1 r. 901 (1985) (current statute).
Alaska	1968	By rule of civil procedure. <i>See</i> ALASKA R. CIV. PROC. 43(j) (1968).
Nevada	1971	By statute. <i>See</i> 1971 Nev. Rev. Stat. ch. 402, § 143, <i>codified at</i> NEV. REV. STAT. § 52.015 (1975).
Arkansas	1976	By statute. <i>See</i> ARK. STAT. ANN. § 28-1001, r. 901(b)(3) (sub. nom. Uniform Rules of Evidence), <i>declared invalid, Ricarte v. State</i> , 290 Ark. 100 (1986)).
North Dakota	1977	By Rule of Evidence. <i>See</i> N.D.R. EVID. 901(b)(2)-(b)(3) (1977).
Wyoming	--	No precedents or sources.