

Critical Issues Affecting the Reliability and Admissibility of Handwriting Identification Opinion Evidence—How They Have Been Addressed (or Not) Since the 2009 NAS Report, and How They Should Be Addressed Going Forward: A Document Examiner Tells All

*Andrew Sulner**

I. INTRODUCTION	634
II. THE NEED TO REPLACE THE SUPPOSED FOUNDATIONAL PRINCIPLES SUPPORTING FORENSIC HANDWRITING EXPERTISE WITH A MORE DEFENSIBLE SET OF PRINCIPLES BASED ON NEUROSCIENCE	638
A. Handwriting Pattern Recognition—An Everyday Occurrence in the Life of the Ordinary Person	640
B. Preliminary Observations on the Classification of Handwriting Forms	641
C. Principles of Human Motor Control Theory as the Foundational Basis for Forensic Handwriting Expertise	642
1. Complexity Theory as Applied to Handwriting	642
D. Establishing Reliability by Demonstrating a Common Sense Justification for the Particular Task Performed by the Forensic Handwriting Expert	644
E. Using Less Absolutist Propositions in Support of Forensic Handwriting Expertise	646
III. LAWYERS AND JUDGES NEED TO KNOW (AND PRACTITIONERS NEED TO OBEY) THE ESSENTIAL REQUIREMENTS FOR PERFORMING RELIABLE FORENSIC HANDWRITING EXAMINATIONS	647
A. Suitability for Comparison and Presence of Sufficient	

* Principal Owner, Forensic Document Examinations, LLC, New York, New York, M.S. (Forensic Science), J.D., George Washington University, Diplomate and immediate past President, Board of Forensic Document Examiners, and past Chair, Jurisprudence Section, American Academy of Forensic Sciences. I would like to thank Michael Risinger for his editorial and other assistance in the writing of this Article, and Edward Imwinkelried for helpful comments on the draft.

Discriminating Writing Features	647
B. Adequacy of Exemplars Used for Comparison Purposes.....	648
1. Contemporaneousness	648
2. Representative (Randomly Selected) Exemplars—Not Self-Serving Exemplars “Cherry Picked” by a Writer Disavowing Authorship	649
3. Obtaining and Examining Handwriting Exemplars from All Persons of Interest in a Limited Pool of Plausible Writers	649
IV. LAWYERS AND JUDGES NEED TO BE MORE COGNIZANT OF THE POTENTIAL IMPACT OF COGNITIVE BIAS ON THE OPINIONS AND TESTIMONY OF FORENSIC HANDWRITING EXPERTS.....	656
A. Sources of Cognitive Bias That Can Unduly Influence the Outcome of Forensic Handwriting Examinations	656
B. Observer Effects: How Examiner Bias Can Unduly Influence Forensic Handwriting Expertise	659
1. Selective Exposure: Choosing Which Evidence to Examine	659
2. Selective Scrutiny: Selectively Evaluating Evidence in a Manner That Favors a Particular Outcome	659
3. Overlooking Differences in Writing Features Due to Observer Effects from the Evidence Itself (“Familiarity Heuristic”).....	660
4. Selective Stopping (“Rush to Judgment” Mindset) ..	662
5. Selective Reevaluation of Evidence and/or Revision of Findings	663
C. Recommendations (Proposed Solutions) for Minimizing Examiner Bias in Forensic Handwriting Investigations ..	663
1. Considering the “Oppositional Hypothesis” First.....	665
2. Considering Alternative Possibilities and Hypotheses (Playing the Role of “Devil’s Advocate”)	666
3. Using the Fischhof Method to Compare Text-Based Writings That Appear Very Similar and as a “Self-Review” of One’s Initial Observations.....	666
4. Using Institutional Debiasing Techniques: Context Management Protocols and Procedures	667
i. Separating the Crime Laboratory or Evidence Analysis Function from the Police and Prosecutorial Functions	667
ii. Using Sequential Unmasking Procedures and Case Managers.....	667
iii. Using Exemplar (Evidence) Lineups and Blind	

2018]

A DOCUMENT EXAMINER TELLS ALL

633

Evidence Submission Protocols	668
V. CASEWORK PEER REVIEWS: THE NEED TO KNOW WHAT THEY ARE AND AREN'T AND WHY DIRECT TESTIMONY ABOUT ANY REVIEW PERFORMED BY A NON-TESTIFYING EXPERT IS INADMISSIBLE AS A MATTER OF LAW	671
A. Types of Casework Peer Reviews	671
1. Administrative Review	671
2. Technical Review	671
3. Verification	672
B. To Verify or Not to Verify?—Should the “V” in “ACE-V” Be a Prerequisite to the Admissibility of Handwriting Identification Opinion Evidence?	674
C. Testimony About Peer Review(s) Performed by a Non- Testifying Expert Constitutes Improper Bolstering and/or Inadmissible Hearsay	677
VI. GATEKEEPING: TRIAL JUDGES AND LAWYERS ARE STILL LARGELY INEFFECTIVE IN PREVENTING THE ADMISSIBILITY OF UNRELIABLE OR EXAGGERATED EXPERT TESTIMONY	685
A. The Responsibility of Gatekeepers Regarding Proffered Expert Testimony	686
B. Factors Relevant and Material to Properly Assessing the Reliability of Proffered Opinion Evidence Under Rule 702	687
C. Court Decisions Illustrating Unreliable and Critically Flawed Testimony on the Part of Presumptively Well- Qualified Handwriting Experts That Should Have Been Excluded by the Trial Judge	688
1. Improper Methodology—Using Exemplar Signatures of a Remote Date to Support an Opinion of Forgery	688
2. Illogical (Unfounded) Conclusion—Unjustifiably Extrapolating from an Accepted Premise to an Unfounded Conclusion	691
i. Incomparability of Writing Features: The Significance of the Questioned Hand Printing Being Unnatural Stick Printing and the Exemplars Being Natural Printing	694
ii. Distorting an Inconclusive Opinion in a Manner That Favors a Particular Outcome	695
iii. Misinterpreting Evidence or Providing Exaggerated Testimony to Support a Favored Outcome	697
3. Improper Methodology—Relying Solely Upon <i>Post</i> <i>Litem Motem</i> Signatures Supplied by A Disclaiming Signatory to Support an Opinion of Forgery (the	

<i>Almeciga case)</i>	698
VII. CONSIDERATIONS AND SUGGESTIONS FOR PROCEEDING FURTHER ALONG “THE PATH FORWARD”	700
A. What the Community of FDE Practitioners Needs to Do	700
1. Restate the Foundational Principles Supporting the Discipline of Forensic Handwriting Analysis as a Technical Skill	700
2. Develop More Robust and Transparent Best Practice Standards for Examining Handwritten Items and Expressing Opinions in Reports and Testimony	701
i. Handwriting Examinations: The Need for Standards to Incorporate Practical Measures Designed to Reduce the Potential for Cognitive Bias	703
ii. Handwriting Examinations: The Need to Delineate the Decision-Making Stages Involved in the Examination and Evaluation Process	707
iii. Handwriting Identification Opinions: The Need to Establish More Meaningful and Appropriate Language for Expressing Qualitative Levels of Confidence (Certitude)	709
iv. Support for and Cooperation with Properly Designed Black Box Studies	714
B. Lawyers and Gatekeepers Need to Be Better Informed and More Diligent in Scrutinizing an Expert’s Methodology	715
1. The Trial Lawyer	715
2. The Trial Judge	716
CONCLUSION	717

I. INTRODUCTION

Although the discipline of forensic document examination requires specialized knowledge with respect to paper, ink, printing methods, detecting alterations and page substitutions, deciphering indented writing, and other subject matters relevant to determining the authenticity or source of a given document, the bulk of a document examiner’s work involves comparison of handwriting of various kinds. This Article will focus only on the handwriting comparison aspect of document examination since that is what has drawn the most attention and criticism with respect to issues of validity and reliability.

In writing this Article, it was not lost on me that the person whose career and contributions are being celebrated in this Symposium is more knowledgeable about forensic handwriting analysis than any other non-practitioner and, regrettably, more knowledgeable than far too many practitioners of this discipline. Many of Professor Michael Risinger's contributions to the relevant literature are routinely cited by courts as authoritative references with respect to seminal issues pertaining to the validity and reliability of forensic handwriting analysis.¹ I consider myself fortunate and privileged to be able to count Michael Risinger as a friend, and even more fortunate to have benefited from his wisdom and insight over many, many years. His viewpoints and critical commentary, more than anything else, have shaped, and in many instances confirmed, my own views as a practicing forensic document examiner (FDE).

By way of introduction, I am a third-generation FDE in private practice for more than forty years, who has also experienced the trials and tribulations of serving the public as a state prosecutor. My education and training in the forensic sciences and the law was comprehensive and multifaceted. It taught me how to think critically when evaluating data or evidence and to consider alternative possibilities, and it reinforced the importance of maintaining neutrality when embarking upon any truth-seeking endeavor. In my chosen field of forensic science, I was fortunate to receive my principal training from an internationally recognized expert in forensic document examination,

¹ See, e.g., D. MICHAEL RISINGER, *Handwriting Identification*, in MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY (David L. Faigman et al., eds., 2d ed. 2002); D. Michael Risinger, *Appendix: Cases Involving the Reliability of Handwriting Identification Expertise Since the Decision in Daubert*, 43 TULSA L. REV. 477 (2007); D. Michael Risinger, *Goodbye to All That, or, a Fool's Errand, by One of the Fools: How I Stopped Worrying About Court Responses to Handwriting Identification (and "Forensic Science" in General) and Learned to Love Misinterpretations of Kumho Tire v. Carmichael*, 43 TULSA L. REV. 447 (2007); D. Michael Risinger, *The Irrelevance, and Central Relevance, of the Boundary Between Science and Non-Science in the Evaluation of Expert Witness Reliability*, 52 VILL. L. REV. 679 (2004); D. Michael Risinger & Mark P. Denbeaux, *Kumho Tire and Expert Reliability: How the Question You Ask Gives the Answer You Get*, 34 SETON HALL L. REV. 15 (2003); D. Michael Risinger, Michael J. Saks, William C. Thompson & Robert Rosenthal, *The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*, 90 CAL. L. REV. 1 (2002); D. Michael Risinger, *Defining the "Task at Hand": Non-Science Forensic Science After Kumho Tire Co. v. Carmichael*, 57 WASH. & LEE L. REV. 767 (2000); D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock*, 64 ALB. L. REV. 99 (2000); D. Michael Risinger, Mark P. Denbeaux & Michael J. Saks, *Brave New "Post-Daubert World"—A Reply to Professor Moenssens*, 29 SETON HALL L. REV. 405 (1998); D. Michael Risinger & Michael J. Saks, *Science and Nonscience in the Courtroom: Daubert Meets Handwriting Identification Expertise*, 82 IOWA L. REV. 21 (1996); D. Michael Risinger, Mark P. Denbeaux & Michael J. Saks, *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification "Expertise,"* 137 U. PA. L. REV. 731 (1989).

Hanna F. Sulner, who happened to be my mother.² I received supplemental training from a retired FBI document examiner and numerous other document examiners and forensic scientists trained in the public and private sectors in various continuing education programs over the years.

Over the last twenty-five years, I have been actively involved in efforts to develop meaningful consensus standards establishing best practices and protocols for FDEs, especially with respect to the examination and comparison of handwritten items, and how the findings and opinions derived from such an examination should be expressed in reports and testimony. I have also been active in exposing flawed and disingenuous testimony on the part of presumptively well-qualified FDEs and in seeking to develop practical measures and safeguards that would enhance the ability of FDEs, lawyers and judges to minimize, if not eliminate, inaccurate and unreliable opinion evidence from forensic handwriting analysis.

This Article endeavors to improve awareness and enhance the legal and judicial responses to critically flawed evidence from handwriting experts by providing lawyers and trial judges with guidelines for determining when evidence from even presumptively well-qualified handwriting experts is unreliable and should be excluded.

² Hanna F. Sulner received her training as a document examiner in Europe. From the age of 16, she studied at the elbow of her father, Professor Julius Fischhof, a pioneer in handwriting analysis who settled in Budapest, Hungary after World War I and won a reputation as Eastern Europe's foremost expert on questioned documents. She studied Criminology in Budapest and in Germany and received a special degree qualifying her to teach the subject of Questioned Documents at the University of Budapest law school. Taking over her father's work after his death in 1944, she quickly inherited his reputation as a meticulous professional as well as his position as official handwriting and document expert to Hungary's courts, police, and military, which she held until February 6, 1949, when she and her husband, Laszlo, escaped to Austria with the aid of American and British intelligence agencies. Four days later, they surfaced in Vienna and achieved international notoriety by denouncing the February 1949 trial of Cardinal Joseph Mindszenty on treason and other charges as a farce, and by displaying microfilms of documents which they had smuggled out of Hungary and which revealed the Communist Government's machinations to discredit and frame the Cardinal by forging his signatures on a purported confession and other incriminating documents. After immigrating to the United States, she promptly resumed her career as an examiner of questioned documents. From 1950 until her death in 1999, Hanna F. Sulner worked in New York as one of the nation's leading authorities on disputed documents, testifying in civil and criminal cases in various state and federal courts throughout the United States and elsewhere, rarely for the losing side. See Robert McG. Thomas Jr., *Hanna F. Sulner, 81, Expert Drawn Into Mindszenty Plot*, N.Y. TIMES (Jan. 19, 1999), <http://www.nytimes.com/1999/01/19/us/hanna-f-sulner-81-expert-drawn-into-mindszenty-plot.html?mcubz=3>. Her professional publications appeared in scientific and legal journals in Europe as well as in such noted American legal periodicals as the *American Bar Association Journal* and *Criminal Law Review*. She authored a well-recognized 400-page reference book, *DISPUTED DOCUMENTS: NEW METHODS FOR EXAMINING QUESTIONED DOCUMENTS* (1966), and *Forged, Altered and Substituted Medical Records*, a comprehensive article that appeared in the October 1971 issue of *Trauma*, an authoritative Medico-Legal journal published by Matthew Bender & Co., New York.

Part II of this Article discusses the foundational tenets that have traditionally been espoused to support the discipline of forensic handwriting analysis, and why there is a need to replace those tenets with less absolutist statements derived from principles of neuroscience and human motor control theory.

Part III provides an overview of the essential requirements for performing reliable forensic handwriting examinations, which lawyers and judges need to know and practitioners should adhere to. A reported opinion from a case in which I was involved is used to illustrate the danger of tunnel vision in any forensic handwriting investigation (focusing only on the “target” of the investigation).

Part IV seeks to illuminate the impact of cognitive bias on forensic handwriting analysis by discussing the biasing influences that can improperly taint and sway an examiner’s decision-making process and render unreliable his or her opinions and testimony, and by providing an overview of the various methods and techniques available for minimizing, if not eliminating, such adverse influences.

Part V examines the three different types of forensic casework peer reviews (administrative review, technical review, and verification); it addresses the relevance of verification in evaluating the reliability of opinion evidence, and discusses recent case law explaining why testimony concerning any type of “review” performed by a non-testifying expert constitutes improper bolstering³ and/or inadmissible hearsay.

Part VI discusses the gatekeeping obligation imposed upon trial judges to screen and exclude unreliable evidence proffered by handwriting experts, and what lawyers need to do, but all too often fail to do, to facilitate the judge’s role as gatekeeper. Several appellate court decisions are reviewed to illustrate instances of unreliable handwriting identification opinion evidence that escaped recognition and successful challenge at the trial level, as well as one recent district court case in which the proffered testimony of a handwriting expert was properly excluded by the gatekeeper.

³ Generally, a party may not bolster or support the credibility of its witness until that credibility has been affirmatively impeached by the opposing party, and even then, it can only do so with evidence that does not violate the hearsay rule. See *generally* EDWARD J. IMWINKELRIED, PAUL C. GIANNELLI, FRANCIS A. GILLIGAN, FREDRIC I. LEDERER & LIESA RICHTER, *COURTROOM CRIMINAL EVIDENCE* § 702 (6th ed. 2016). As discussed in Part V.C, *infra*, several court decisions have held that a witness’s testimony about casework peer review by someone not testifying is inadmissible on the ground of “improper bolstering” of the credibility of a witness on direct examination in advance of any attack on the witness’s credibility. However, I am grateful to Professor Edward Imwinkelried for pointing out that the more appropriate legal basis for excluding such attempted corroboration testimony would be hearsay outside of any hearsay exception.

Finally, Part VII concludes with my thoughts and perspectives on what can and should be done to further enhance the reliability of handwriting identification opinion evidence that is presented in civil and criminal cases. This Part also discusses what can and should be done to remedy the inequality of resources in criminal cases that effectively deprives defendants and their counsel of the ability to acquire the expertise needed to properly expose flaws and weaknesses in opinion evidence proffered by prosecution handwriting experts.

II. THE NEED TO REPLACE THE SUPPOSED FOUNDATIONAL PRINCIPLES SUPPORTING FORENSIC HANDWRITING EXPERTISE WITH A MORE DEFENSIBLE SET OF PRINCIPLES BASED ON NEUROSCIENCE

Forensic document examiners (FDEs) have traditionally premised the claim of scientific validity and reliability of handwriting identification on two asserted principles or tenets: (1) handwriting is unique, meaning that no two people write exactly alike (the principle of Uniqueness or *Inter-writer* Variability); and (2) no person can produce an exact duplicate of his or her signature or write exactly the same way twice (the principle of *Intra-writer* Variability).⁴

Given the strong form in which these two tenets are usually expressed, their absolute truth is by no means obvious (nor is it necessary for the process of handwriting analysis to yield probative and reliable results under many circumstances, as we will see). But even assuming their truth, why does this establish the credentials of the process of handwriting identification as a science?

I have always maintained, contrary to the position espoused by most FDEs, that handwriting identification is not a science, but a technical skill which, if performed properly, can have great probative value, for reasons I will go into later in this Article.⁵ Such views have long made me an outsider to the “guild” identified and described in 1997 by Professor Risinger and his

⁴ See, e.g., Diane Harrison, Ted M. Burkes & Danielle P. Sieger, *Handwriting Examination: Meeting the Challenges of Science and the Law*, FORENSIC SCI. COMMNC'NS (Oct. 2009), https://www.fbi.gov/about-us/lab/forensic-science-communications/fsc/oct2009/review/2009_10_review02.htm/. These two tenets are sometimes supplemented by a third, that no person can exceed his or her writing skill level at a given point in time (the principle of Writer Skill Threshold Limitation), a principle which is virtually tautological, self-proving, and requires no validation. However, it does not apply in the factual context of most cases, and does leave open the question of how to determine the skill level of both the questioned writing and the candidates for writership. *Id.*

⁵ When I so testified before U.S. District Court Judge Miriam Cedarbaum in *Boulé v. Hutton*, 138 F. Supp. 2d 491 (S.D.N.Y. 2001), Judge Cedarbaum interrupted my testimony to comment that during her career on the bench, she had heard testimony from many handwriting experts, and I was the first one to testify that handwriting identification was not a “science.”

coauthors.⁶ Now, however, my longstanding position that absolutist statements concerning uniqueness and intra-writer variability are as yet unproven, and likely unprovable, appears to be gaining more respectability as academic scientists examine the foundations of the traditional pattern-matching forensic disciplines and come to broadly similar conclusions. In addition, recent federal district court decisions excluding or limiting the proffered testimony of handwriting experts in criminal and civil cases have likewise concluded that regardless of whether handwriting analysis is characterized as science or technical skill, there has been inadequate testing and insufficient data to support the scientific validity of the two fundamental principles (Uniqueness and *Intra-writer Variability*) that have forever been espoused by virtually all FDEs as the underlying basis for handwriting identification.

For instance, in the recent case *United States v. Kelly*,⁷ the handwriting at issue involved cursive handwriting. In limiting the proffered expert testimony to a discussion of the expert's methodology and the similarities and differences he observed in handwriting features, the court noted:

In his testimony, Mr. Shiver stated that in order to definitively prove that no individual writes exactly like any other, one would have to collect the handwriting of every person. Mr. Shiver also stated that collecting such data is practically impossible. *The inability to prove this tenet with certainty does not in and of itself*

⁶ Professor Risinger and his coauthors described the guild mentality in handwriting expertise circles in D. Michael Risinger, Michael J. Saks & Mark P. Denbeaux, *Brave New "Post-Daubert World: A Reply to Professor Moenssens*, 29 SETON HALL L. REV. 405, 408 (1998) (commenting on Andre A. Moenssens, *Handwriting Identification Evidence in the Post-Daubert World*, 66 UMKC L. REV. 251 (1997)):

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

Id.

⁷ *United States v. Kelly*, No. 4:15-cr-01632 (D. Ariz. 2015).

render Mr. Shiver's identification opinion unreliable. However, the two studies Mr. Shiver states support this tenet do not provide a sufficient basis for his opinion. One study leads to the conclusion that within a sample of 1500 writers, very few write similarly; the other leads to the conclusion that handwriting is not determined by a person's genetics. Without more, this Court cannot conclude from these studies that there is sufficient data to support an identification opinion.⁸

Similarly, in *United States v. Johnsted*,⁹ the proffered hand printing identification testimony of Gale Bolsover, a U.S. Postal Service FDE, was excluded in its entirety. The court cited the following reasons for excluding the testimony:

- The studies cited by the Government and its expert fail to support the principles of Uniqueness (*Inter-writer Variability*) and *Intra-Writer Variability* that are claimed as the foundational underpinnings of handwriting analysis.¹⁰
- Bolsover conceded that she knew of no studies supporting the second foundational principle that no individual writes exactly the same way twice.¹¹

A. *Handwriting Pattern Recognition—An Everyday Occurrence in the Life of the Ordinary Person*

While there has been some progress in accepting that the absolutist statements historically claimed as the foundational tenets of forensic handwriting analysis are not capable of empirical proof, there is still some resistance within the FDE community to abandoning these absolutist statements in favor of more modest, and more defensible, claims concerning the diagnosticity of handwriting comparison in regard to various different tasks which are presented in actual practice, based on a more nuanced understanding of such claims.

Handwriting recognition is a phenomenon that ordinary people are invariably familiar with and routinely encounter in their daily lives, which renders it unlike any other forensic pattern recognition phenomenon. Common experience confirms that under some circumstances, we can

⁸ *Id.* at 4–5 (emphasis added).

⁹ 30 F. Supp. 3d 814 (W.D. Wis. 2013).

¹⁰ *Id.* at 822.

¹¹ *Id.* at 818.

recognize a writer by means of visual cues appearing in his or her writing. Usually, such writing belongs to a relative, friend or co-worker whose writing pattern has been unconsciously stored and imbedded in the visual memory center of our brain after numerous viewings, and is thereafter subject to automatic recall upon viewing it again. Presumably, for this recognition to be achieved, the brain is making decisions based on features that pictorially characterize the writing. For lay persons, this process is relatively straightforward as only a limited number of pictorial memories are in play and an incorrect judgment may have no adverse consequences. For forensic handwriting experts, every sample of handwriting presented, be it questioned or known, presents a set of handwriting features that are unfamiliar and have not been stored in the observer's visual memory, which therefore requires methodical examinations of the sets of questioned and known handwriting independently to ascertain the features that characterize each set.

B. *Preliminary Observations on the Classification of Handwriting Forms*

There have been many systems proposed for the classification of handwriting, some quite elaborate, and often with different purposes in mind. These are well surveyed by Roy A. Huber and A. M. Headrick in their treatise *Handwriting Identification: Facts and Fundamentals* (1999).¹² For this paper, it is best to stick to fundamentals. Signatures are properly classified separately because they represent a special class of writing that is personal to the writer in many physical and emotional dimensions, and thus presents both special opportunities and special problems in determining whether a putative signature by a specified person was made by that person (is authentic¹³). Signatures may be text-based, stylized (symbolic); or mixed text-stylized. Non-signature generation of written letters has been described by Huber and Headrick as falling into four principle categories: (1) Cursive writing, in which letters are connected; (2) "Manuscript" writing, in which letters are disconnected and are designed similar to upper and lower case printing characters; (3) Hand lettering, sometimes called block printing or block lettering, in which letters are separately structured and often designed as upper case letters; and (4) Composites usually combining cursive writing and "manuscript" writing, with some letters connected and others not.¹⁴ It is

¹² ROY A. HUBER & A.M. HEADRICK, *HANDWRITING IDENTIFICATION: FACTS AND FUNDAMENTALS* 152–60 (1999).

¹³ "Authenticity" in this context is a specialized type of writership attribution in which a signature is attributed to the person whose name or alleged stylization appears on the document. A signature not attributed to that putative author is not authentic, or inauthentic.

¹⁴ HUBER & HEADRICK, *supra* note 12, at 95–101.

important to note here that it is quite common for a person's ordinary hand to be manuscript writing, and quite uncommon for it to be hand lettering. Both of these, not being connected cursive, are often misleadingly lumped together as "hand printing" even though they present different sub-tasks from the standpoint of determining the sufficiency of features for purposes of writer attribution.

C. *Principles of Human Motor Control Theory as the Foundational Basis for Forensic Handwriting Expertise*

Handwriting is an over-learned skill that becomes automatic and habitual (it's practiced until the writer "can't get it wrong"). For example, the image of your signature and the specific movement sequence required to produce that image are imbedded in your brain's motor control memory centers. Hence, if you were to sign your signature in the air with your forefinger or foot acting as the writing instrument, your forefinger or foot would follow the same movement sequence that is followed when signing your name on paper with pen in hand. Since handwriting is a behavioral artifact of the human motor control process, once you understand that process you can draw reasonable inferences from examining its byproduct.

1. Complexity Theory as Applied to Handwriting

Numerous human motor control studies conducted by neuroscientists have confirmed the difficulty of repeating or simulating a series of rapidly connected movements ("Complexity Theory").¹⁵ As applied to handwriting, the following principles can be reliably stated:

- A. The more complex a handwriting sample is, the more difficult it becomes for others to simulate it without leaving behind tell-tale indicia of simulation (e.g., line tremor, unnatural pen lifts, gaps or hesitations, spatial and proportional disturbances).

¹⁵ See, e.g., Jeffrey J. Brault & Rejean Plamondon, *Complexity Measure of Handwritten Curves: Modeling of Dynamic Signature Forgery*, 23 IEEE TRANSACTIONS ON SYS., MAN, & CYBERNETICS 400 (1993); Henry S.R. Kao, Daniel T.L. Shek & Elbert S.P. Lee, *Control Modes and Task Complexity in Tracing and Handwriting Performance*, 54 ACTA PSYCHOLOGICA 69 (1983); Ruud G.J. Meulenbroek, & Gerard P. van Galen, *Perceptual-Motor Complexity of Printed and Cursive Letters*, 58 J. EXPERIMENTAL EDUC. 95 (1990); Rulldolf E. van der Platts & Gerard P. van Galen, *Effects of Spatial and Motor Demands in Handwriting*, 22 J. MOTOR BEHAV. 361 (1990); Gerard P. van Galen, *Structural Complexity of Motor Patterns: A Study on Reaction Times and Movement Times of Handwritten Letters*, 46 PSYCHOL. RES. 49 (1984); Alan M. Wing, *Response Timing in Handwriting*, in INFORMATION PROCESSING IN MOTOR CONTROL AND LEARNING (George E. Stelmach ed., 1978).

2018]

A DOCUMENT EXAMINER TELLS ALL

643

- B. The more complex a handwriting sample, the more likely it will contain features that deviate from other writers.¹⁶
- C. The greater the number of times the pen is required to change direction, the longer the line length over which turning points occur, and the greater the overall speed of execution, the more complex the visual image appears.

Empirical support exists for the above stated complexity theory relationships, and the assessment of the complexity of handwriting has been reported in related fields of research.¹⁷ Research studies have validated the ability of FHEs to evaluate the parameters that contribute to the perceived complexity of handwriting.¹⁸

Examples of Signatures of Varying Complexity

Three signatures of varying complexity (simple, moderately complex, and very complex) are displayed below:



¹⁶ This relationship is related to the theory of class and individual characteristics that has traditionally been offered as an underpinning for forensic handwriting identification.

¹⁷ Linda C. Alewijnse, C. Elisa van den Heuvel & Reinoud D. Stoel, *Analysis of Signature Complexity*, 21 J. FORENSIC DOCUMENT EXAMINATION 37 (2011); See generally MICHAEL P. CALIGURI & LINTON A. MOHAMMED, *THE NEUROSCIENCE OF HANDWRITING: APPLICATIONS FOR FORENSIC DOCUMENT EXAMINATION* (2012).

¹⁸ Bryan J. Found, Douglas Rogers & Robert Schmittat, *Recovering Dynamic Information From Static Handwriting Traces Using Angular Differential Software*, 6 J. QUESTIONED DOCUMENT EXAMINATION 17 (1997); Huub J.J. Hardy, *Dynamics of the Writing Movement: Physical Modelling and Practical Applications*, 5 J. FORENSIC DOCUMENT EXAMINATION 1 (1992); Tahnee N. Dewhurst, Bryan J. Found & Douglas Rogers, *The Relationship Between Quantitatively Modelled Signature Complexity Levels and Forensic Document Examiners' Qualitative Opinions on Casework*, 18 J. FORENSIC DOCUMENT EXAMINATION 21 (2007); Bryan J. Found & Douglas Rogers, *The Forensic Investigation of Signature Complexity*, in *HANDWRITING AND DRAWING RESEARCH: BASIC AND APPLIED ISSUES* 483–92 (Marvin L. Simner et al. eds., 1996); Emily J. Will, *Pilot Study: Inferring Relative Speed of Writing From the Static Trace*, 22 J. FORENSIC DOCUMENT EXAMINATION 55 (2012).

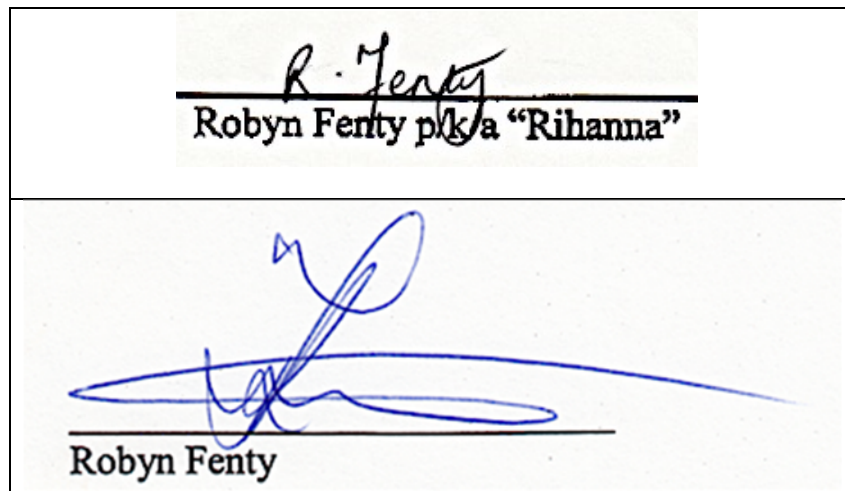


Figure 1. Signatures of Varying Complexity

The bottom two signatures in Figure 1 above belong to the star entertainer, Rihanna. The first was signed early in her career, the second after she became an international celebrity. Recognizing the inherent value of her signature on memorabilia, like many other sports and entertainment celebrities, Rihanna intentionally developed a highly complex signature to make it more difficult for would-be forgers to successfully simulate her signature.

D. Establishing Reliability by Demonstrating a Common Sense Justification for the Particular Task Performed by the Forensic Handwriting Expert

As discussed above, the reliability of forensic handwriting analysis cannot be predicated upon the uniqueness of handwriting. Rather, it must be predicated upon (a) the perceptual ability of FDEs to observe and detect subtle similarities and differences between two sets of writings; and (b) their diagnostic ability to make reasonable inferences from such an analysis regarding the likelihood that the two sets of writings were written by the same or different writers.

The standard for admissibility of proffered expert testimony varies somewhat from jurisdiction to jurisdiction. For the purposes of this Article, I will concentrate on admissibility criteria under Federal Rule of Evidence 702. For expert testimony to be admissible under FRE 702, the trial judge must have sufficient reason to believe that (a) the expert possesses sufficient skill in regard to performing a particular task at hand, and (b) reliable inferences can be drawn from the reliable performance of that particular task. Under FRE 702, as under *Daubert* and *Kumho*, trial courts must scrutinize

not only the principles and methods used by the expert, but also whether those principles and methods have been properly applied to the facts of the case. *Daubert's* list of specific factors neither necessarily nor exclusively applies to all experts or in every case,¹⁹ and *Kumho* acknowledges that whatever the claim of expertise may be, when you're dealing with skill claims, the belief warrant need not necessarily come from empirical testing.²⁰ A demonstrable critical common sense justification for a warranted belief in reliability supplants the need for independent black box or statistical validation studies, even though such studies are always preferable.

The fact that the ordinary person is very familiar with handwriting and can easily understand the mechanism that produces it lessens the necessity for *formal* validation of a handwriting expert's ability to make accurate decisions regarding signature authenticity or handwriting source attributions.²¹ Once you understand the mechanism by which a phenomenon is produced, you can readily see when an opinion based upon that mechanism is sufficiently reliable to be admitted in evidence under a common sense approach. Critical common sense, the nature of the evidence being analyzed and characterized, and the ability of the ordinary person to comprehend and independently assess the expert's characterization of that evidence provides a sufficient belief warrant for the validity of the characterization. For example, demonstrable evidence of shared characteristics between two sets of handwriting can provide a strong basis for a reasonable person to find that both sets of writings were made by the same person beyond a reasonable doubt.²²

Kumho Tire makes it clear that under the mandate of *Daubert*, evaluation must be directed to the reliability of the expertise in the specific "task at hand" at issue in the case.²³ Let's assume that the task at hand requires distinguishing natural writing tremor from faked (artificial) tremor in a questioned (disputed) signature. Let us further assume that the tremor observed in the stroke pattern of the known signatures of the subject writer

¹⁹ *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 141 (1999).

²⁰ *Id.* at 150.

²¹ See *United States v. Starzecpyzel*, 880 F. Supp. 1027, 1044 (S.D.N.Y. 1995) ("The ability of juries to perform the crucial visual comparisons relied upon by handwriting experts cuts against the danger of undue prejudice from the mystique attached to 'experts.'").

²² "A reasonable doubt is a fair doubt based on reason, logic, common sense, or experience. It is a doubt that an ordinary reasonable person has after carefully weighing all of the evidence, and is a doubt of the sort that would cause him or her to hesitate to act in matters of importance in his or her own life. It may arise from the evidence, or from the lack of evidence, or from the nature of the evidence." U.S. COURT OF APPEALS FOR THE THIRD CIRCUIT, MODEL CRIMINAL JURY INSTRUCTIONS § 3.06 (Apr. 2015), [http://www.ca3.uscourts.gov/sites/ca3/files/2012 Chapter 3 Rev.pdf](http://www.ca3.uscourts.gov/sites/ca3/files/2012%20Chapter%203%20Rev.pdf).

²³ *Kumho Tire*, 526 U.S. at 141. See analysis in Risinger, *Defining the "Task at Hand,"* *supra* note 1.

reflects wiggly and irregular strokes indicating the writer's inability to control the writing instrument, whereas the tremor observed in the stroke pattern of the questioned signature reflects a series of concatenated sharply angled up and down strokes that look like shark's teeth, indicating extreme control over the writing instrument. In this instance, the neuroscience of motor control provides an understanding of the mechanism that produces handwriting and a common sense justification for the inference that the questioned signature was written by someone who possessed motor control rather than someone who lacked it.

E. *Using Less Absolutist Propositions in Support of Forensic Handwriting Expertise*

As discussed above, the FDE community needs to abandon the theory of uniqueness as the underpinning for forensic handwriting expertise and replace it with a neuroscience model based upon human motor control behavior and complexity theory. More modest fundamental principles should be used as support for handwriting expertise, such as the following:

- (1) Given an adequate amount of skillfully executed, complex (non-simplistic) writing, the likelihood that handwriting by different writers will be distinguishable from each other is far greater than the likelihood that handwriting by different writers will be indistinguishable from each other.²⁴
- (2) The smaller the source population of possible writers, the greater the likelihood that a specific writer can be accurately identified as the source of the questioned writing.

²⁴ In discussing the neurobiological principles underlying handwriting variation, and the principle that the more complex the writing, the harder it should be to simulate, and the less chance of mistaking it for another individual's handwriting, the late Bryan Found, a pioneering Australian forensic document examiner (FDE) whose views were very much in line with mine, likewise asserted that "under normal conditions, given a sufficient amount of writings, no two skilled writers are likely to produce handwritten images that are exactly the same in terms of the combination of construction, line quality, formation variation and text structure features." Bryan J. Found & Douglas Rogers, *A Consideration of the Theoretical Basis of Forensic Handwriting Examination: The Application of "Complexity Theory" to Understanding the Basis of Handwriting Identification*, 4 INT'L. J. FORENSIC DOCUMENT EXAMINERS 109 (1998).

III. LAWYERS AND JUDGES NEED TO KNOW (AND PRACTITIONERS NEED TO OBEY) THE ESSENTIAL REQUIREMENTS FOR PERFORMING RELIABLE FORENSIC HANDWRITING EXAMINATIONS

A. *Suitability for Comparison and Presence of Sufficient Discriminating Writing Features*

Handwriting is an acquired skill involving a multifaceted perceptual-motor task in which neuromuscular commands direct a very complex set of coordinated movements of the fingers, wrist, elbow and shoulder. The discriminating features of writing include elements of style (e.g., letter formations, spatial and proportional relationships between letters and words, and formatting features) and execution (e.g., speed and fluidity of writing movements). It is the totality (combination) of the discriminating, habitual writing habits that forensic document examiners compare and evaluate in cases involving handwriting identification and/or signature verification. Essentially, the pictorial, structural and line quality features that are perceived to characterize two sets of writing specimens are independently assessed and then compared *inter se* to determine whether they are sufficiently similar to support the hypothesis of common authorship or sufficiently dissimilar to support the hypothesis that different writers produced the two sets of writings.

At the outset of any handwriting investigation, the examiner makes value judgments about whether the questioned writing is sufficiently devoid of distortion or disguise to render it suitable for comparison purposes, and whether it contains enough distinguishing features to support a decision regarding source attribution.²⁵ Such judgments are discretionary and examiner dependent.

Once the questioned writing is adjudged to be suitable for comparison purposes, the examiner then evaluates the quantity and quality of the exemplar (known) writing to assess its adequacy for comparison purposes.

Let us consider signatures first. To determine whether a questioned signature is genuine, a trained forensic handwriting expert focuses on the intricate details that make up the component (structural) parts of the signature and the relative speed and fluency (rhythm) with which those details are executed. An attempt to duplicate the signature of another person based upon a known sample or “model” of that person’s signature is referred to as a *forgery by simulation* or *simulated forgery*. In so doing, the forger attempts to duplicate the normal and natural writing habits and abilities of

²⁵ In the case of the first (top) signature displayed in Figure 1, *supra*, it is so simplistic that a would-be forger would have little difficulty producing a simulation that would escape detection as a forgery. Hence, any conclusion concerning the authenticity of such a simplistic signature should be considered unreliable and of no probative value.

another while simultaneously discarding his or her own customary writing habits and abilities.

A skilled signature, be it text-based, stylized, or a combination of both, is the product of a series of concatenated curvilinear strokes. The number of times the writing instrument is required to change direction and the relative orientation of the curvilinear strokes within a given signature establish its complexity. To imitate the precise writing movements so as to accurately replicate the combination of curvilinear strokes found within a complex signature of another person and to do so with the same relative speed of execution and fluency found within that other person's genuine signature(s) is a task that is extremely difficult if not impossible to achieve. Hence, the greater the complexity of the handwritten image, the more difficult it becomes to successfully simulate it. Essentially, the difficulty encountered in copying complex movements is the cornerstone of forensic signature analysis.

B. *Adequacy of Exemplars Used for Comparison Purposes*

Obtaining an adequate number of samples of an individual's normal writing is an essential requirement in investigating whether or not such individual authored a questioned or disputed handwritten item; these samples are termed *exemplars*. The exemplars must be sufficient in quantity to provide a sound basis for evaluating and ascertaining the natural range of variation found within the subject individual's handwriting or signature pattern. Variations found within the same person's writing or signature pattern are often referred to as "*intra-writer*" differences, whereas "*inter-writer*" differences refer to *dissimilarities* that are attributable to another writer. In any case involving questioned writings or disputed signatures, the critical task for the forensic document examiner is to ascertain whether apparent differences are *intra-writer* differences indicative of common authorship, or *inter-writer* differences evidencing different writers. The significance of the essential requirements of this task will be highlighted in one or more of the illustrative cases discussed later in this Article.

1. Contemporaneousness

Perhaps the most important factor in assembling good exemplars is contemporaneousness. Ideally, the exemplars should be written as close as possible to the alleged date(s) of preparation of the questioned writing(s).²⁶

²⁶ The ideal exemplars for handwriting comparisons are contemporaneous normal course of business writings that conform to the same format and context as the questioned material, i.e., prepared in the same writing style (cursive or printed), and containing the same characters, letters, or words. The textual format of the exemplar writing needs to be similar to that encountered in the questioned writing so that the examiner is "comparing apples to apples."

Contemporaneity of the exemplar writing is particularly critical in those instances where the examiner bases his or her opinion upon dissimilarities observed between the questioned and exemplar writing. Best practices require contemporaneous exemplars for comparison purposes because they tend to be far more representative of the subject's writing habits and skill at the time the questioned item was purportedly written; as the time gap between the exemplars and the questioned writing becomes greater, the exemplars have the tendency to be less representative and more unreliable. Hence, it should come as no surprise that a principal source of error in disputed signature cases is when the handwriting expert bases an opinion of forgery on exemplar signatures of a remote date or on an inadequate amount of exemplar signatures.

2. Representative (Randomly Selected) Exemplars—Not Self-Serving Exemplars “Cherry Picked” by a Writer Disavowing Authorship

Best practices also require experts to reduce the likelihood that the exemplar writing submitted to them for comparison purposes is self-serving and not representative of the full range of writing features attributed to a specific writer. For example, in a disputed signature case, if the process by which exemplar signatures are selected permits a disclaiming signatory to “cherry pick” the exemplars to be supplied for comparison purposes, the selection process itself is tainted and inherently unreliable. The selection process must allow for the random selection of exemplars to reduce the likelihood that exemplars were selected with the intent to provide spurious support for an unmeritorious claim of forgery. In many signature comparison cases, upon obtaining a truly representative sampling of the disclaiming party's signature pattern, what at first glance were perceived as “apparent differences” are oftentimes demonstrated and proven to be “normal variations” within the same person's signature pattern (*intra*-writer differences), and hence *prima facie* proof of genuineness.

3. Obtaining and Examining Handwriting Exemplars from All Persons of Interest in a Limited Pool of Plausible Writers

One of the cardinal sins committed in a forensic handwriting investigation is to perform one-to-one comparisons between the handwriting appearing in a questioned item and only the exemplar writing attributable to the “suspect” or “target of the investigation,” especially when there exists a limited pool of plausible writers. As a general principle, every effort should

The ideal exemplars for signature comparisons are signatures executed on a date contemporaneous with the date of the questioned or disputed signature, and preferably on documents of similar import and under comparable circumstances of execution.

be made to keep the examiner “blinded” from learning the identity of the suspect in a handwriting investigation as such contextual information can unconsciously influence the perception and decision-making of the handwriting expert, as described in Part IV of this Article. The need to examine writing samples from others in the “web” of plausible writers should be self-evident: it facilitates keeping the examiner blinded to the “suspect” or person of greatest interest; it enables the examiner to see if other writers display writing features similar to those observed in the questioned writing(s); and it can serve to eliminate individuals as the writer of the questioned writing. This is particularly important in cases involving hand printing, which may not be as distinctive as cursive writing.

Early in my career, shortly after I stopped working as a state prosecutor for the Queens County District Attorney’s Office, I was hired as a forensic document examiner in what proved to be a case that attracted a lot of local media coverage and reinforced my conviction that FDEs involved in forensic handwriting investigations should always be required to examine writing samples from other plausible writers or persons of potential interest, and, to the extent possible, be precluded from receiving any case information that is irrelevant to the handwriting comparison task they are being asked to perform until the task is completed.

The case involved a high profile disciplinary hearing in which the Westchester County Division of Public Safety charged Police Officer William P. Shaughnessy with serious misconduct that could have led to job termination as well as future criminal charges. John D. Ryan, a former New York State prosecutor, was officially appointed as Special Hearing Officer to hear and evaluate the evidence and report his findings and determination. It was alleged that between February and December of 1981, Shaughnessy was part of a conspiracy in which approximately two hundred (200) mail subscription forms were fraudulently completed in the names of several police officers and their spouses, resulting in hundreds of books and magazines being sent to their homes and offices. Two of Shaughnessy’s superior officers, Lt. Stephen Fischer and Sgt. James Fleming, received 95 and 78 such unsolicited mailings, apparently in retribution for their involvement and cooperation in the so-called “cooping investigation,” a clandestine NYPD Internal Affairs investigation of police officers sleeping on the job, which resulted in charges being brought against twenty-two (22) police officers and a great deal of unwanted media attention.

Shaughnessy was charged with the intent to defraud, harass, annoy, and alarm others by falsely and fraudulently completing approximately ten of these mail subscription order forms, eight in the name of Stephen Fisher and two in the name of James Fleming. Far more serious was the allegation that Shaughnessy and his co-conspirators cut the lug-nuts on Lieutenant

Fleming's personal automobile, which resulted in a near fatal accident on April 3, 1981, when the left rear tire fell off while Fleming was driving on a highway. Fleming testified that at around that same time, while at the office of their same personal attorney on separate business, Fleming and Shaughnessy had a conversation in which Shaughnessy expressed an opinion that the mailings would never stop. Fleming was involved in the cancellation of over 200 unsolicited mailings as well as the procedure utilized with respect to forwarding evidence to the postal authority's crime laboratory. He enjoyed a friendly working relationship with Terry Loftus, the Postal Inspector assigned to investigate the unsolicited mailings, having worked with him for five years on stolen welfare check cases. Fleming was also aware that Police Officer Robert Duncan was separately indicted in connection with the cooping investigation and had pled guilty.

Shaughnessy, a recipient of the Bronze Medal of Honor for military service in Viet Nam, was not indicted or even considered a suspect in the cooping investigation. He adamantly denied writing, in whole or in part, any mail subscription forms, all of which contained alphanumerical hand printing. There was no eyewitness account of Shaughnessy's involvement in any of the activities with which he was charged, and no admissions of any kind were made by him. As noted in the published decision of Special Hearing Officer Ryan, "[t]his case turns on a proper analysis of the expert testimony submitted by each party,"²⁷ and "the testimony of each handwriting expert was . . . , without a doubt, the most significant and probative part of this case."²⁸

The County's handwriting expert was Carl J. Raichle, a presumptively well-qualified FDE employed by the U.S. Postal Crime Lab in New York City.²⁹ Mr. Raichle's initial report, dated February 8, 1982, was based upon his examination and comparison of the hand printing appearing in the ten questioned mail subscription forms that Shaughnessy was charged with writing and the known hand printing of Shaughnessy appearing in his memo book and personnel folder. In this report, Mr. Raichle concluded that Shaughnessy wrote seven out of the ten subscription forms at issue and that

²⁷ *In re Guido v. Shaughnessy*, 1983 Extra LEXIS 3, at *28 (Westchester Cty., N.Y. Dep't Pub. Safety Nov. 18, 1983).

²⁸ *Id.* at *36.

²⁹ Mr. Raichle earned a B.S. in Police Science from John Jay College of Criminal Justice, trained four years in the New York City Police Department Crime Lab's Questioned Documents Unit, attended FBI and U.S. Secret Service specialized courses, and was employed as a forensic document examiner by the New York City Police Department Crime Lab for four years and the U.S. Postal Crime Lab for six years. He was certified by the American Board of Forensic Document Examiners (ABFDE) and testified approximately thirty times in state and federal court. *Id.* at *15–16.

“it is conceivable that Shaughnessy wrote [the other three].”³⁰

Thereafter, Mr. Raichle requested and received request (demand) exemplars that Shaughnessy provided to Postal Inspector Loftus following Mr. Raichle’s first report. No other writer’s exemplars were submitted to or requested by Mr. Raichle. Based upon these additional Shaughnessy exemplars, Mr. Raichle issued his second report, dated June 10, 1982, which now concluded that Shaughnessy wrote each of the ten mail subscription forms he was charged with writing.

Mr. Raichle had also received a packet from Lt. James Fleming that contained samples of P.O. Garret Morrison’s hand printing. Fleming submitted these additional hand printing exemplars to Mr. Raichle because Shaughnessy had claimed that his and Morrison’s hand printing were similar.³¹ Mr. Raichle acknowledged that he never compared Morrison’s hand printing to the hand printing on the questioned mail subscription forms because Shaughnessy had already been identified as the author of the forms.³²

Nearly a year after issuing his second report, Mr. Raichle received for analysis 10 additional mail subscription forms that Shaughnessy was not formally charged with writing. This resulted in Mr. Raichle issuing a third report, dated June 21, 1983, in which he concluded that at least four different writers completed these ten additional mail subscription forms and that Shaughnessy wrote four of them.

I was formally retained as a handwriting expert on behalf of Respondent Shaughnessy in late December 1982. From the very start, I insisted on obtaining exemplar hand printing from all the police officers in Shaughnessy’s unit, including supervisory personnel, and that I be supplied with all the mail subscription forms analyzed by Mr. Raichle. The fact that Mr. Raichle never insisted on examining writing exemplars from writers other than the charged party (Shaughnessy) was shocking to me given that standard protocol (best practices) in any anonymous handwriting investigation requires doing so. Sadly, thirty-five years later, this case still serves as an excellent learning tool for experts, lawyers and judges because it illustrates the importance of examining handwriting exemplars from all

³⁰ *Id.* at *16.

³¹ During the investigative phase of the case, Police Officer Garrett T. Morrison, who had access to the records and reports of various police officers, was shown each of the ten original subscription forms at issue and denied authorship of each. At trial, Morrison appeared under subpoena as a witness for Respondent Shaughnessy. When offered the opportunity by Respondent’s attorney to provide handwriting exemplars, Morrison refused to do so without a court order upon the advice of counsel. *Id.* at *20.

³² This is a prime example of *target shifting* and *selective stopping*. Morrison’s handwriting was not compared to the questioned handwriting because the target, Shaughnessy, was already identified.

2018]

A DOCUMENT EXAMINER TELLS ALL

653

persons of interest when dealing with a limited population of plausible writers and the dangers inherent in examining only those written by the prime suspect in the investigation. This case also demonstrates the need to erect a “Chinese wall” between the expert analyzing the evidence and the investigator handling the case.

My analysis of the hand printing at issue revealed numerous significant dissimilarities between the hand printing in question and the known hand printing of Shaughnessy. These differences were readily observable when reviewing either the post litem motem request exemplars provided by Shaughnessy or his normal course of business exemplars that pre-dated the case (ante litem motem exemplars). My opinion and testimony was that there was absolutely no evidence that Shaughnessy wrote any of the ten subscription forms he was charged with writing or that he wrote any of the ten uncharged subscription forms, four of which Mr. Raichle attributed to Shaughnessy. I further testified and demonstrated that four of the ten charged forms were written by Patrick Duncan (who had already been criminally convicted for his involvement), another two appeared to have been written by Duncan, and the remaining four contained significant similarities to the hand printing of Police Officer Garrett Morrison, and were, beyond all doubt, not written by Shaughnessy.³³

³³ “Mr. Sulner, due to lack of sufficient available exemplars from Morrison, would not positively identify Morrison as the author of these items. However, he did testify that due to the number and probative weight to be assigned to the numerous similarities, any indication of authorship is on Morrison. Shaughnessy, he testified, is, beyond all doubt, not the author.” *In re Guido v. Shaughnessy*, 1983 Extra LEXIS 3, at *25 (citations omitted).

The chart below indicates which of the salient features of the questioned hand printing were found by Mr. Raichle to be similar or dissimilar to the hand printing features of Shaughnessy (the only writer whose hand printing he compared).

HW Features in Questioned Documents (Mail Subscription Forms)	SHAUGHNESSY Pre-Existing Hand Printing	SHAUGHNESSY Request (Demand) Hand Printing	MORRISON Hand Printing	DUNCAN Hand Printing
"eR" in "FischeR" (misplaced upper case "R")	YES	YES	<i>Not Examined</i>	<i>Not Examined</i>
"eR" in "FischeR" (pen drag between "e" and "R")	YES	YES	<i>Not Examined</i>	<i>Not Examined</i>
"L" in "PoLice" (misplaced upper case "L" and its relative height)	YES	YES	<i>Not Examined</i>	<i>Not Examined</i>
"e-s" combination - "e" is always HIGHER than "s"	Always LOWER	Always LOWER	<i>Not Examined</i>	<i>Not Examined</i>
"p-h" combination - "p" is always HIGHER than "h"	Always LEVEL	Always LEVEL	<i>Not Examined</i>	<i>Not Examined</i>
"G" (formed like numeral "6" vs. traditional form)	NEVER	NEVER	<i>Not Examined</i>	<i>Not Examined</i>
"J" (form - always without horizontal top)	NEVER	NEVER	<i>Not Examined</i>	<i>Not Examined</i>
"M" (form - middle leg always hits baseline)	NEVER	NEVER	<i>Not Examined</i>	<i>Not Examined</i>
"N" (form - right side always curved inward)	NEVER	NEVER	<i>Not Examined</i>	<i>Not Examined</i>
"S" (bowl vertically aligned)	NEVER	NEVER	<i>Not Examined</i>	<i>Not Examined</i>
zeros shorter in zip codes ("1oo22")	NEVER	NEVER	<i>Not Examined</i>	<i>Not Examined</i>
"5"s (top always skewed, not level)	Always LEVEL	Always LEVEL	<i>Not Examined</i>	<i>Not Examined</i>
i-dot pattern (dot always placed high above and to far right of stem)	Always CLOSE to stem	Always CLOSE to stem	<i>Not Examined</i>	<i>Not Examined</i>
variant case forms of "t"s in middle or end of word ("eaTing" or "eaTing")	NEVER	NEVER	<i>Not Examined</i>	<i>Not Examined</i>
"U"s and "u"s NEVER end with a tail	Always a TAIL	Always a TAIL	<i>Not Examined</i>	<i>Not Examined</i>

2018]

A DOCUMENT EXAMINER TELLS ALL

655

The chart below reflects my findings regarding the distribution across three writers (Shaughnessy, Morrisown and Duncan) of the same silent hand printing features.

HW Features in Questioned Documents (Mail Subscription Forms)	SHAUGHNESSY Pre-Existing Hand Printing	SHAUGHNESSY Request (Demand) Hand Printing	MORRISON Pre-Existing Hand Printing	DUNCAN Pre-Existing Hand Printing
"eR" in "FischeR" (misplaced upper case "R")	YES	YES		
"eR" in "FischeR" (pen drag between "e" and "R")	YES	YES		
"L" in "PoLice" (misplaced upper case "L" and its relative height)	YES	YES	YES	
"e-s" combination - "e" is always HIGHER than "s"	Always LOWER	Always LOWER		
"p-h" combination - "p" is always HIGHER than "h"	Always LEVEL	Always LEVEL		
"G" (formed like numeral "6" vs. traditional form)	NEVER	NEVER	YES	
"J" (form - always without horizontal top)	NEVER	NEVER	YES	YES
"M" (form - middle leg always hits baseline)	NEVER	NEVER	YES	
"N" (form - right side always curved inward)	NEVER	NEVER	YES	YES
"S" (bowl vertically aligned)	NEVER	NEVER		
zeros shorter in zip codes ("10022")	NEVER	NEVER	YES	YES
"5"s (top always skewed, not level)	Always LEVEL	Always LEVEL	YES	YES
i-dot pattern (dot always placed high above and to far right of stem)	Always CLOSE to stem	Always CLOSE to stem	YES	
variant case forms of "t"s in middle or end of word ("eaTing" or "eaTing")	NEVER	NEVER		
"U"s and "u"s NEVER end with a tail	Always a TAIL	Always a TAIL		

The results obtained from my examination of exemplar writings from writers other than just the target (suspect) made it clear that more of the salient features were found in the writings of Morrison and Duncan than in Shaughnessy's.³⁴ This, coupled with my testimony as to the significant

³⁴ The reader is referred to the decision *In re Guido v. Shaughnessy*, 1983 Extra LEXIS 3, (Westchester Cty., N.Y. Dep't Pub. Safety Nov. 18, 1983), for a detailed analysis of the

dissimilarities between Shaughnessy's hand printing and the questioned hand printing, is what enabled the trier of fact to reject the testimony of the government's expert and conclude:

I must concur with the findings of Mr. Sulner. Aside from the similarities in the "L" in "Police," there do not appear to be any significant similarities in the questioned writings and those of P.O. Shaughnessy. The numerous distinctions between the two (2) sets of writings compel the conclusion that Shaughnessy was not the author of the questioned writings.³⁵

IV. LAWYERS AND JUDGES NEED TO BE MORE COGNIZANT OF THE POTENTIAL IMPACT OF COGNITIVE BIAS ON THE OPINIONS AND TESTIMONY OF FORENSIC HANDWRITING EXPERTS³⁶

A. *Sources of Cognitive Bias That Can Unduly Influence the Outcome of Forensic Handwriting Examinations*

Contextual bias is the most common form of cognitive bias encountered in forensics. It occurs when potentially biasing background information that is irrelevant to the discipline-specific task assigned to an examiner (e.g., examining and comparing handwriting) is conveyed to the examiner before the examiner has completed the task and reached a conclusion. It is not uncommon for forensic document examiners to be "briefed" about the background of the case surrounding the document(s) being submitted to them for forensic handwriting analysis.³⁷ Such extraneous information usually suggests the outcome preferred or desired by the party requesting the analysis, and consequently, has the potential to unduly influence and distort the examiner's visual perception and evaluation of the handwriting evidence submitted. In the law enforcement or criminal justice setting, potentially biasing information usually concerns the crime itself, the criminal background of the suspect, or knowledge of a confession or some other form of physical or testimonial evidence linking the suspect to the crime.

handwriting evidence presented in this case, particularly with respect to the handwriting features that the government's handwriting expert mischaracterized as "significant" similarities attributable to Shaughnessy.

³⁵ *In re Guido v. Shaughnessy*, 1983 Extra LEXIS 3, at *47–48 (Westchester Cty., N.Y. Dep't Pub. Safety Nov. 18, 1983).

³⁶ The material in Part IV originally appeared in Andrew Sulner, *Handwriting: Cognitive Bias*, in WILEY ENCYCLOPEDIA OF FORENSIC SCIENCES 1–15 (Allan Jamieson & Andre Moenssens eds., 2014).

³⁷ Larry S. Miller, *Bias Among Forensic Document Examiners: A Need For Procedural Changes*, 12 J. POLICE SCI. & ADMIN. 407 (1984).

The idea that document examiners should be insulated from all information about an investigation except necessary, domain-specific information is not novel. William E. Hagan's 1894 treatise on the examination of disputed handwriting and signatures contained the following commentary highlighting the need to keep document examiners "blinded" from such biasing influences:

[T]he examiner must depend wholly upon what is seen [in the forensic examination], leaving out of consideration all suggestions or hints from interested parties; and if possible it best subserves the conditions of fair examination that the expert should not know the interest which the party employing him to make the examination has in the result. Where the expert has no knowledge of the moral evidence or aspects of the case in which signatures are a matter of contest, there is nothing to mislead him, or to influence the forming of an opinion; and while knowing of the case as presented by one side of the context might or might not shade the opinion formulated, yet it is better that the latter be based entirely on what the writing itself shows, and nothing else.³⁸

Motivational bias on the part of forensic experts can be attributed to a variety of factors, and research on motivated reasoning has shown that an individual's reasoning processes are more readily biased when the individual is motivated by goals other than accuracy.³⁹ Wharton provides the following commentary regarding motivational bias on the part of handwriting experts:

It is well known that in cases of peculiar difficulty, when the difference, if there be any, between two handwritings is only noticeable by perceptions, the most sensitive experts, no matter how conscientious, often take unconsciously such a bias from the party employing them as to give to their judgment the almost infinitely slight impulse that turns the scale; nor is it strange that, in an instrument so delicate, aberrations from its true course should be produced by attractions or repulsions otherwise unappreciable. If an expert could be absolutely secluded from such extraneous influences, his judgment might be depended on at least for impartiality. This, however, is impracticable. A jury is bound, therefore, to accept the opinion of an expert as to handwriting, even when uncontradicted, as an argument rather

³⁸ WILLIAM E. HAGAN, A TREATISE ON DISPUTED HANDWRITING AND THE DETERMINATION OF GENUINE FROM FORGED SIGNATURES 82 (1894).

³⁹ See DAN SIMON, IN DOUBT: THE PSYCHOLOGY OF THE CRIMINAL JUSTICE PROCESS 25 (2012) (citing Ziva Kunda, *The Case for Motivated Reasoning*, 108 PSYCHOL. BULLETIN 480 (1990)).

than a proof; and to make allowance for all the disturbing influences by which the judgment of the expert may be moved.⁴⁰

Wharton's view of when the judgments of handwriting experts are most vulnerable to bias is confirmed by psychological research indicating that forensic practitioners are less likely to be swayed by potentially biasing influences when the evidence is clear-cut and unambiguous.⁴¹ Simply put, it is far more difficult to rationalize a desired outcome in the face of very strong, if not irrefutable, evidence to the contrary.

Contextual and motivational influences can produce confirmation bias, the tendency to seek out and interpret evidence in ways that support or confirm pre-existing beliefs and desires. Conflicts between truth-seeking goals and outcome-oriented goals are often fueled by the adversarial nature of the legal process itself.⁴² The 2009 NAS Report⁴³ and research studies⁴⁴ indicate that forensic practitioners assigned to evaluate evidence may be motivated to see their side of a case prevail, which can lead them to endorse a biased view of the evidence that is consistent with their adoption of an adversarial, outcome-oriented role instead of an objective truth-seeking one. Moreover, "tough-on-crime" attitudes prevalent within the law enforcement community tend to foster confirmation biases that leave prosecutors, investigators, and forensic specialists in crime laboratories more inclined to prioritize the value of obtaining a conviction of the accused over the countervailing priority of protecting the accused from a wrongful conviction.⁴⁵ These potentially biasing influences render handwriting and

⁴⁰ FRANCIS WHARTON, A COMMENTARY ON THE LAW OF EVIDENCE IN CIVIL ISSUES 711–12 (3d ed. 1888).

⁴¹ See SIMON, *supra* note 39, at 25 (citing Karl Ask, Anna Rebelius & Pär Anders Granhag, *The "Elasticity" of Criminal Evidence: A Moderator of Investigative Bias*, 22 APPLIED COGNITIVE PSYCHOL., 1245 (2008); Itiel E. Dror & David Charlton, *Why Experts Make Errors*, 56 J. FORENSIC IDENTIFICATION 600 (2006)).

⁴² See generally Jennifer L. Mnookin, *Expert Evidence, Partisanship, and Epistemic Competence*, 73 BROOK. L. REV. 1009 (2008).

⁴³ NAT'L RES. COUNCIL OF THE NAT'L ACADS., STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 122–24 (2009), <https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf> [hereinafter 2009 NAS REPORT].

⁴⁴ See SIMON, *supra* note 39, at 26 (citing Dan Simon, Doug Stenstrom & Stephen J. Read, *On the Objectivity of Investigations: An Experiment* (Aug. 18, 2008), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1130103).

⁴⁵ See, e.g., Stephen B. Perrot & Donald M. Taylor, *Attitudinal Differences Between Police Constables and Their Supervisors: Potential Influences of Personality, Work Environment, and Occupational Role*, 22 CRIM. JUST. & BEHAV. 326, 333–37 (1995) (explaining that levels of authoritarianism views differ between constables and supervisors); Richard K. Wortley & Ross J. Homel, *Police Prejudice as a Function of Training and Outgroup Contact: A Longitudinal Investigation*, 19 LAW & HUM. BEHAV. 305 (1995) (explaining that police prejudices vary according to policing experiences); Karl Ask & Pär Anders Granhag, *Motivational Sources of Confirmation Bias in Criminal Investigations: The Need for Cognitive Closure*, 2 J. INVESTIGATIVE PSYCHOL. & OFFENDER PROFILING 43 (2005).

2018]

A DOCUMENT EXAMINER TELLS ALL

659

other forensic experts vulnerable to making erroneous decisions about the evidence they evaluate.

Absent the implementation of practical bias minimizing procedures when evidence is submitted to and evaluated by handwriting experts in civil or criminal cases, it is unlikely to expect such experts to be kept “blinded” from domain-irrelevant information or other potential biasing influences. Some of the recommended changes that have been proposed for minimizing bias in handwriting investigations are discussed below in Part V.C.

B. Observer Effects: How Examiner Bias Can Unduly Influence Forensic Handwriting Expertise

Observer effects refer to the ways in which an examiner’s perception and interpretation of evidence can be influenced by the examiner’s preconceived beliefs and motives, or by the surrounding context, which can include background information conveyed to the examiner as well as the evidence itself, the latter being a phenomenon often overlooked. Examiner bias in forensic handwriting investigations can influence how examinations and comparisons are performed, the visual perceptions and observations of the examiner, the findings and opinions drawn from evaluating and comparing questioned and known writings, and the manner in which the examiner testifies in court. Some of the mechanisms and mental processes by which such cognitive bias can operate are discussed below.

1. Selective Exposure: Choosing Which Evidence to Examine

Observer effects refer to the ways in which an examiner’s perception and interpretation of evidence can be influenced by the examiner’s preconceived beliefs and motives, or by the surrounding context, which can include background information conveyed to the examiner as well as the evidence itself—the latter being a phenomenon often overlooked. Examiner bias in forensic handwriting investigations can influence how examinations and comparisons are performed, the visual perceptions and observations of the examiner, the findings and opinions drawn from evaluating and comparing questioned and known writings, and the manner in which the examiner testifies in court. Some of the mechanisms and mental processes by which such cognitive bias can operate are discussed below.

2. Selective Scrutiny: Selectively Evaluating Evidence in a Manner That Favors a Particular Outcome

A handwriting expert’s selective scrutiny of evidence occurs when the expert searches only for evidence that will confirm the expert’s favored outcome. An example would be where a handwriting expert’s favored outcome is common authorship and in the course of examining the evidence, the expert’s attention is disproportionately focused on looking for

similarities in writing features between the questioned and exemplar writing, thereby failing to meaningfully search for or recognize the presence of differences in writing features between the two sets of writings. As discussed in Part IV.B.3, the Fischhof Method of upside-down writing comparison developed by my grandfather in the early 1900s can help to minimize the risk that differences between two sets of writings will be overlooked in cases of particular difficulty or ambiguity.⁴⁶

3. Overlooking Differences in Writing Features Due to Observer Effects from the Evidence Itself (“Familiarity Heuristic”)

Much of the literature in the emerging field of cognitive forensics has focused on observer effects resulting from extraneous (domain-irrelevant) contextual information, neglecting the effect the evidence itself can have on the observer’s visual perceptions of that evidence. Although cognitive psychologists have long been aware that familiarity can cause oversight of unusual events or situations, this heuristic (which I have labeled the “familiarity heuristic”) has been largely overlooked as a factor contributing to observer effects in forensic handwriting examinations and comparisons.⁴⁷

Julius Fischhof, a pioneer in the field of questioned documents and Eastern Europe’s leading handwriting expert in the early twentieth century,⁴⁸ recognized that in the context of text-based handwriting, the familiarity of letters or words can unconsciously contribute to an examiner’s failure to observe or recognize salient writing features, most notably differences between two sets of text-based signatures or handwritten items that appear to be very similar. Fischhof discovered that by comparing such questioned and known signatures or handwriting upside down, the examiner is not subconsciously influenced by reading individual letters or words and has a more objective view of writing features.⁴⁹ In essence, the Fischhof Method

⁴⁶ The FBI Laboratory recently revised its Standard Operating Procedures (SOPs) and training materials for Examining Friction Ridge Impressions. “The training materials provide techniques for avoiding bias, such as tracing the entire unknown print before looking at the known, reconstructing the deposition pressure to ‘reverse engineer’ distortions, or *turning the print upside down or looking at it obliquely to get a different perspective*.” U.S. DEP’T OF JUST. OFF. OF THE INSPECTOR GEN., A REVIEW OF THE FBI’S PROGRESS IN RESPONDING TO THE RECOMMENDATIONS IN THE OFFICE OF THE INSPECTOR GENERAL REPORT ON THE FINGERPRINT MISIDENTIFICATION IN THE BRANDON MAYFIELD CASE 28 (June 2011), <https://oig.justice.gov/special/s1105.pdf> (emphasis added).

⁴⁷ D. Michael Risinger et al., *The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*, 90 CAL. L. REV. 1, 15 (2002) (“The flexibility of the human cognitive system permits us to ‘tune’ ourselves to perceive some things and ignore other things, usually so automatically and seamlessly that we rarely realize we are doing it. This tuning process results in ‘selective attention’ to information.”).

⁴⁸ See Thomas Jr., *supra* note 2.

⁴⁹ JULIUS FISCHHOF, NEW METHOD OF COMPARING HANDWRITING (1927); HANNA F.

of upside-down comparison offers the examiner a means of avoiding undesirable observer effects from the very thing being observed—the handwriting—by preventing the ocular distraction resulting from following written characters or words readily familiar to the observer. It serves to minimize the cognitive noise and interference resulting from the familiarity heuristic associated with observations of text-based handwriting by altering the handwritten image into something that is unrecognizable, thereby tricking the brain into thinking it is seeing an unfamiliar image. By providing a totally different visual perspective of the very same evidence, the Fischhof Method affords FDEs the type of visual feedback that helps them avoid overlooking differences in relevant writing features that might impact the accuracy of their judgments about handwriting. It is akin to the professional proofreader's strategy of reading material backward when checking for misspellings in order to avoid becoming distracted by content issues.

In *In re Last*, a 1989 New York State Surrogate Court decision involving conflicting expert testimony about the validity of a signature appearing on a shareholders agreement between two brothers,⁵⁰ the Court commented favorably on the Fischhof Method:

The petitioner's expert, a well-known authority and author in the field of handwriting analysis, concluded that the signature of Walter Last on the shareholder's agreement was a forgery. Her testimony included a detailed analysis of the subject signature with a comparison to known exemplars of the decedent's signature. She employed an "upside-down" technique in which a known and a questioned signature are compared after they are inverted. Since there is a natural tendency to read words instead of noting variations in characters, this method allegedly gives the examiner a truer basis for comparison. Employing photographic enlargements of known signatures and the questioned signature, and acknowledging that no two signatures of the same person are exactly alike, she emphasized differences in both primary and secondary characteristics and opined that the questioned signature was not that of the decedent.

There was an attempt to show, both by testimony that the Last brothers signed each other's signature, and by noting certain characteristics in Bert's signature, that the questioned signature "is more identical to the characteristics in Bert Last's handwriting . . . than with Walter Last's signature." The court, after being advised that the petitioner did not intend to show that

SULNER, DISPUTED DOCUMENTS: NEW METHODS FOR EXAMINING QUESTIONED DOCUMENTS 48 (1966).

⁵⁰ *In re Last*, No. 1983-1177, 1989 N.Y. Misc. LEXIS 933 (N.Y. Sur. Ct. Jan. 9, 1989).

Bert had committed the forgery, ruled the testimony irrelevant and barred further questioning along these lines.

The respondent countered with another expert, a trained examiner of questioned documents. He described his method of examining the questioned signature and comparing it with a series of known signatures of the decedent. The expert considered such features as skill, slant, speed, spacing proportions, relative size, and upper case letter versus lower-case comparisons. Pen stops, hesitations, tremors and possible tracing were also taken into account. Pictorial aspects and design forms were reviewed, particularly as they applied to variation (no two signatures of a person are exactly alike). On the basis of these tests, this expert concluded that the questioned signature is that of the decedent. When asked why the questioned signature appeared to have a break between two letters, he said the lack of a “connecting stroke” was insignificant, attributing it to a normal variation. Under extensive cross-examination, he explained apparent inconsistencies in the signatures, such as hooks, straight lines and spaces. He found all fell within the parameters of variation contemplated in multiple, one-author signatures.

The expert testimony offered by the petitioner, while lacking in certain respects, was more convincing than that presented by the respondent. *The analysis conducted by the petitioner’s expert, particularly the “upside-down” comparison, was credible and persuasive.* The explanation offered by the respondent’s expert was insufficient to eliminate glaring differences between the signatures, particularly as regards spacing. The normal variation present in everyone’s signature does not account for the divergence in primary characteristics, as cogently explained by the petitioner’s expert.⁵¹

4. Selective Stopping (“Rush to Judgment” Mindset)

Selective stopping occurs when an investigation prematurely terminates further inquiries after having found some evidence to support a favored outcome, but before adequate consideration was given to alternative hypotheses or the existence and availability of evidence tending to refute the favored outcome. This “rush to judgment” mindset, a byproduct of confirmation bias, has contributed to flawed FBI investigations in several

⁵¹ *Id.* at *5–7 (emphasis added).

high-profile cases, such as the wrongful arrests and subsequent exonerations of Richard Jewell in connection with the 1996 bombing of Atlanta's Centennial Olympic Park that killed 1 individual and injured 117 others,⁵² and Brandon Mayfield in connection with the 2004 Madrid train bombings that killed 191 individuals and wounded 1,800.⁵³

5. Selective Reevaluation of Evidence and/or Revision of Findings

More often than not, domain-irrelevant background information about a given case is conveyed to an examiner at the time the evidence is initially submitted for analysis. Sometimes, the information is obtained afterward, as when the examiner learns that his/her findings are inconsistent with test results obtained from forensic analysis of other items of evidence in the case, or from analysis of the very same evidence by a different analyst. Cross-communication of findings from analysis of the same or other evidence can unduly influence objectivity of handwriting experts, and it has been raised as a possible source of error in many cases involving handwriting identifications, especially where the disclosure of such information has prompted the examiner to refine or change the initial conclusion after "reevaluating" the very same evidence. In response to the 2009 NAS Report and revelations that cognitive bias contributed to laboratory and practitioner errors in some high-profile criminal cases,⁵⁴ the FBI laboratory has reportedly discontinued its long-standing practice of allowing forensic examiners in one discipline unit to know the findings of forensic examiners in another discipline unit and to confer with one another in the event of conflicting results.

C. *Recommendations (Proposed Solutions) for Minimizing Examiner Bias in Forensic Handwriting Investigations*

Forensic document examiners and others in the forensic science community have historically dismissed cognitive bias as a factor contributing to examiner errors in casework, insisting that such errors are caused by incompetence or dishonesty rather than domain-irrelevant

⁵² See Luke Cyphers, *Richard Jewell Keeps Being Remembered as a Suspect, Not the Hero of the 1996 Atlanta Olympics Bombing*, N.Y. DAILY NEWS (Aug. 29, 2016), <http://www.nydailynews.com/news/national/seeking-bomb-shelter-article-1.891938>.

⁵³ See A REVIEW OF THE FBI'S PROGRESS, *supra* note 46. See also Robert B. Stacey, *Report on the Erroneous Fingerprint Individualization in the Madrid Train Bombing Case*, FBI LAB. SERVS. (Jan. 2005), https://archives.fbi.gov/archives/about-us/lab/forensic-science-communications/fsc/jan2005/special_report/2005_special_report.htm; Steven T. Wax & Christopher J. Schatz, *A Multitude of Errors: The Brandon Mayfield Case*, CHAMPION MAG. (Sept./Oct. 2004). The facts of the case and Mayfield's legal claims against the government are fully reported in *Mayfield v. United States*, 504 F. Supp. 2d 1023 (D. Or. 2007).

⁵⁴ 2009 NAS REPORT, *supra* note 43, at 122–24.

contextual or motivational influences. Consequently, there has been a long-standing reluctance on the part of the forensic community at large to acknowledge the need to develop internal procedures and strategies designed to minimize the likelihood of having the objectivity of forensic decision-making compromised by potentially biasing influences. However, there now exists a substantial body of empirical research reported in peer-reviewed scientific and legal journals and to the forensic science community at large that clearly establishes the susceptibility of handwriting, fingerprint, and other pattern recognition experts to having the results of their examinations and comparisons cognitively contaminated and unduly influenced by domain-irrelevant contextual information and motivational bias.⁵⁵ With more and more stakeholders recognizing and understanding the insidious manner in which cognitive contaminants can be toxic to one's neutrality, proposals and recommendations for minimizing examiner bias are now receiving considerably more attention within the forensic science, legal, and academic communities, as reflected in some of the more recent peer-

⁵⁵ See, e.g., Keith Inman & Norah Rudin, *Sequential Unmasking: Minimizing Observer Effects in Forensic Science*, in *ENCYCLOPEDIA OF FORENSIC SCIENCES* 542–48 (Jay A. Siegel & Pekka J. Saukko eds., 2d ed. 2013); SIMON, *supra* note 39; Sulner, *Handwriting: Cognitive Bias*, *supra* note 36; David Canter et al., *Cognitive Bias in Line-up Identifications: The Impact of Administrative Knowledge*, 53 *SCI. & JUST.* 83 (2013); Simon A. Cole, *Implementing Counter-Measures Against Confirmation Bias in Forensic Science*, 2 *J. APPLIED RES. MEMORY & COGNITION* 61 (2013); Itiel E. Dror, *Cognitive Forensics and Experimental Research about Bias in Forensic Casework*, 52 *SCI. & JUST.* 128 (2012); Itiel E. Dror et al., *New Application of Psychology to Law: Improving Forensic Evidence and Expert Witness Contributions*, 2 *J. APPLIED RES. MEMORY & COGNITION* 78 (2013); Itiel E. Dror, *Practical Solutions to Cognitive and Human Factor Challenges in Forensic Science*, 4 *FORENSIC SCI. POL'Y & MGMT* 105 (2013); Bryan Found & John Ganas, *The Management of Domain Irrelevant Context Information in Forensic Handwriting Examination Casework*, 53 *SCI. & JUST.* 154 (2013); Saul M. Kassin, *Why Confessions Trump Innocence*, 67 *AM. PSYCHOLOGIST* 431, 433–34 (2012); Saul M. Kassin et al., *The Forensic Confirmation Bias: Problems, Perspectives, and Proposed Solutions*, 2 *J. APPLIED RES. MEMORY & COGNITION* 42 (2013); Dan E. Krane et al., *Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation*, 53 *J. FORENSIC SCI.* 1006 (2008); Jeff Kukucka & Saul M. Kassin, *Do Confessions Taint Perceptions of Handwriting Evidence? An Empirical Test of the Forensic Confirmation Bias*, 2 *LAW & HUM. BEHAV.* 1 (2013); Risinger, Saks, Thompson & Rosenthal, *supra* note 1, at 1; Reinoud D. Stoel, Itiel E. Dror, & Larry S. Miller, *Bias Among Forensic Document Examiners: Still a Need for Procedural Changes*, 46 *AUSTRALIAN J. FORENSIC SCI.* 91 (2014); D. Michael Risinger, *How to Undermine Proffered Expert Opinion Infected with the Precursors of Bias*, Speech at the 66th Annual Meeting of the American Academy of Forensic Sciences (Feb. 17, 2014); Andrew Sulner, Chair, Bias in Forensics Workshop, & Barry Scheck, Co-Chair, Bias in Forensics Workshop, *Examining the Sources and Impacts of Bias on Perceptual and Cognitive Judgments Made by Forensic Experts, Strategies for Excluding or Impeaching Expert Testimony Tainted by Bias, and Proposed Solutions for Minimizing or Inhibiting Biasing Influences*, Speeches at the 66th Annual Meeting of the American Academy of Forensic Sciences (Feb. 17, 2014); Andrew Sulner, *Examining Sources of Bias and Illustrating Their Impact on Handwriting Opinions and Testimony of Forensic Document Examiners*, Speech at the 66th Annual Meeting of the American Academy of Forensic Sciences (Feb. 17, 2014).

reviewed publications and presentations addressing this topic.⁵⁶ Unfortunately, the FDE community has been slow in incorporating bias control measures into existing standards governing best practices for performing handwriting examinations.

In the case of forensic handwriting investigations, virtually all procedures and strategies that can be used to minimize examiner bias involve either implementing examiner debiasing techniques or restructuring institutional context management and evidence testing protocols, as briefly summarized below.

1. Considering the “Oppositional Hypothesis” First

As domain-irrelevant information invariably enters the scene through the mouths of lawyers or clients intent on convincing the handwriting expert of the merits of their claim(s), healthy skepticism on the part of the expert goes a long way toward ensuring neutrality in the analysis and evaluation of handwriting evidence. In considering the oppositional hypothesis first, an examiner approaches the investigation with the mindset of having been hired by the adverse or oppositional party. In this way, the examiner is forced to consider the least favored hypothesis first and elaborate on the reasons for

⁵⁶ *Id.* See also e.g., D. Michael Risinger, Bias Control: The National Commission on Forensic Science, the National Institute of Standards and Technology (NIST) and the Draft Guidance on Cognitive Bias Effects from the Forensic Science Regulator for England and Wales, Speech at the 67th Annual Meeting of the American Academy of Forensic Sciences (AAFS) (Feb. 17, 2015); Andrew Sulner et al., Human Factors in Forensic Science: Why Cognitive Bias Can Lead to Flawed Expert Opinions and Testimony, How Its Influence Can Be Minimized, and What Challenges Testifying Experts and Judges Can Expect to Face at the Plenary Session, Speeches at the 67th Annual Meeting of the AAFS (Feb. 18, 2015). All speaker presentations are viewable on YouTube. See Dan Simon, *Cognitive and Motivational Causes of Investigative Error*, YOUTUBE (Apr. 22, 2015), <https://www.youtube.com/watch?v=46mS8gvAYMU&feature=youtu.be>; William C. Thompson, *Are Good Intentions and Willpower Sufficient to Overcome Contextual Bias? A Look at the Psychological Literature*, YOUTUBE (Apr. 22, 2015), https://www.youtube.com/watch?v=bzXft_13Ass&feature=youtu.be; Saul Kassin, *Confessions in Context: Why Confessions Corrupt Forensic Perceptions and Judgments*, YOUTUBE (Apr. 24, 2015), <https://www.youtube.com/watch?v=hC7EBZCnWNs&feature=youtu.be>; Andrew Sulner, *Why Forensic Scientists Should Embrace the Concept of Bias Control: A Practitioner's Perspective*, YOUTUBE (Apr. 22, 2015), <https://www.youtube.com/watch?v=QzH2tNVdqfw&feature=youtu.be>; Barry Scheck, *Beginning A Dialogue: Human Factors and Forensic Pathology*, YOUTUBE (Apr. 22, 2015), <https://www.youtube.com/watch?v=5cjbkVUIDOs&feature=youtu.be>; Andrew Baker, History, *Cognitive Bias, Incompetence, and Corruption Are Not the Same Thing*, YOUTUBE (Apr. 22, 2015), <https://www.youtube.com/watch?v=wiePh1COEYE&feature=youtu.be>; Andrew Sulner, Bias Effects in Forensic Handwriting Investigations and Expert Testimony: An Insider's View, Speech at the 67th Annual Meeting of the AAFS (Feb. 17, 2015); Andrew Sulner, Chair, Bias in Forensics Workshop, & Barry Scheck, Co-Chair, Bias in Forensics Workshop, Cognitive Bias Issues in the Forensic Analysis of Pattern and Impression Evidence and in Medicolegal Evaluations, Speeches at the 67th Annual Meeting of the AAFS (Feb. 17, 2015).

rejecting it. Only then does the examiner consider the most favored hypothesis.

2. Considering Alternative Possibilities and Hypotheses
(Playing the Role of “Devil’s Advocate”)

Considering all plausible alternative possibilities before deciding on a particular one is essential to the integrity of any type of investigation. Promoting a “devil’s advocate” mindset in which thinking “outside the box” is encouraged should therefore be prioritized in the training and continuing education of all forensic experts, as contrarian and critical thinking skills are needed in order to be able to both generate and properly evaluate plausible alternative hypotheses. This is particularly important in the case of handwriting, as its physical appearance can be significantly affected by a variety of environmental and motivational factors (for example, awkward writing position, the influence of drug or alcohol intoxication/withdrawal, the import of the document itself, and deliberate attempt at disguise).

It has also been suggested that examiners should not be allowed to summarily dismiss alternative possibilities and hypotheses, and that any refutations should be accompanied by documentation describing in detail the reasons for rejecting each alternative possibility.

3. Using the Fischhof Method to Compare Text-Based
Writings That Appear Very Similar and as a “Self-Review”
of One’s Initial Observations

The Fischhof Method of upside-down comparison described earlier in this Article can be used as a possible safeguard against overlooking differences in salient writing features whenever a handwriting expert is confronted with two sets of text-based writings that appear quite similar. This method of comparison can also afford an examiner a “fresh new look” at the evidence, enabling observations from the initial analysis to be measured against observations derived from inverted image comparisons of the very same evidence. Optimally, such a self-conducted review should take place at a time well after the initial handwriting analysis so as to reduce the likelihood of any “interference” produced by recall of observations made during the initial analysis.

4. Using Institutional Debiasing Techniques: Context Management Protocols and Procedures

i. Separating the Crime Laboratory or Evidence Analysis Function from the Police and Prosecutorial Functions

The 2009 NAS Report recommended separating the crime laboratory function from any law enforcement department or agency, theorizing that a truly autonomous crime laboratory would mitigate, if not remove, the institutional pressures placed on crime laboratory analysts to produce results that favor the police or prosecution's theory of the case, and would foster a more neutral mindset that prioritizes the truth-seeking goal.⁵⁷ To date, little headway has been made in implementing this concept due to practical and economic considerations, as well as institutional resistance to change.

ii. Using Sequential Unmasking Procedures and Case Managers

Context management protocols involve shielding the examiner from domain-irrelevant information and employing "sequential unmasking" procedures to control the order in which domain-relevant but potentially biasing information is "unmasked" and disclosed to the examiner. Ideally, the examiner is kept as blind as possible for as long as possible, and remains unaware of domain irrelevant information until all examinations and tests are completed.

Optimally, a case manager who is privy to all the facts of the case is responsible for determining which evidence to test, and for evaluating and interpreting the test results in the context of the case in order to assess whether the test results support an inculpatory or exculpatory hypothesis. The case manager should also possess, or have access to, relevant subject matter expertise, as difficult decisions may need to be made about what information is domain-relevant and when and how such information should be obtained and disclosed to the examiner.

The strict protocol for sequential unmasking requires that after looking at the questioned item(s) and before looking at any exemplar(s), the examiner must determine (and make a written record of) the specific distinguishing features that the examiner would rely on in deciding whether to associate or disassociate the questioned item(s) with the exemplar(s).⁵⁸ This procedural

⁵⁷ See 2009 NAS REPORT, *supra* note 43, at 23–24 ("Recommendation 4: To improve the scientific bases of forensic science examinations and to maximize independence from or autonomy within the law enforcement community, Congress should authorize and appropriate incentive funds to the National Institute of Forensic Science (NIFS) for allocation to state and local jurisdictions for the purpose of removing all public forensic laboratories and facilities from the administrative control of law enforcement agencies or prosecutors' offices.")

⁵⁸ See Inman & Rudin, *supra* note 55, at 544–45; Krane, *supra* note 55, at 1006; Risinger, Saks, Thompson & Rosenthal, *supra* note 1; William C. Thompson, *Painting the Target*

requirement is deemed a necessary safeguard against target-shifting, in which knowledge about features contained in the exemplar(s) influences the examiner's interpretation of the questioned item(s) and the examiner's decision about which features are relevant and irrelevant for comparison purposes.⁵⁹

Although sequential unmasking procedures can be implemented with relative ease, most, if not all, forensic laboratories in the United States have not done so for handwriting investigations. This may be due to the fact that the method by which a handwriting expert selects the salient writing features to be used for comparison purposes is subjective and examiner-dependent, there being no standardized best-practice protocol for how such feature selection should be made, let alone documented. The Document Examination Unit of the Victoria Police Forensic Services Department in Australia, however, most recently embarked on a pilot study using a modified version of the sequential unmasking protocol for handwriting cases.⁶⁰

iii. Using Exemplar (Evidence) Lineups and Blind Evidence Submission Protocols

In investigations seeking to determine whether a handwritten item can be attributed to a particular source writer, the FDE is often presented with the questioned item(s) and only the suspected (targeted) writer's reference item(s), such as handwriting exemplars. Some commentators from the legal and scientific communities have criticized the suggestiveness inherent in such a procedure, arguing that exemplar lineups should be used for handwriting identifications and other types of evidential source attributions in much the same way that photo lineups are used for eyewitness identifications.⁶¹ As the same deficiencies that make a photo lineup unduly suggestive make an exemplar lineup unduly suggestive, both types of lineups require presenting similar-looking "fillers" ("foils") to the observer (the handwriting expert or the eyewitness). Thus, in blind exemplar lineups, the examiner would receive the handwritten item(s) in question along with an array of similar-looking handwriting exemplars from a group of anonymous

Around the Matching Profile: The Texas Sharpshooter Fallacy in Forensic DNA Interpretation, 8 LAW, PROBABILITY & RISK 257, 274.

⁵⁹ See, e.g., Inman & Rudin, *supra* note 55, at 542–48; Krane, *supra* note 55, at 1006; D. Michael Risinger, The Need for Sequential Unmasking—With Some Thoughts on How to Undermine Proffered Expert Testimony Infected with the Precursors of Bias, Speech at the 66th Annual Meeting of the AAFS (Feb. 17, 2014).

⁶⁰ See Found, *supra* note 55, at 154–58, for a detailed description of the Australian protocol and how their sequential unmasking procedures were modified so that only the essential information required for performing the requisite handwriting examinations and comparisons is available to the examiner.

⁶¹ *Id.*

individuals, including the suspected writer, and the examiner would receive no information or cognitive cues that might unduly influence the examiner to reach a particular outcome. To ensure that the analyst receives no improper cues from the person(s) tasked with submitting the evidence to be analyzed, it has been suggested that exemplar lineups be double-blind, meaning that both the analyst and the individual(s) submitting the evidence or arranging the exemplar lineup(s) not know the identity of the suspect or the preferred outcome.⁶²

In theory, every forensic pattern recognition discipline that requires comparisons between unknown (questioned) items and known reference items (exemplars) in order to determine the source of the unknown item(s) can benefit from the use of exemplar lineups, as the presence of a large number of “fillers” resembling the questioned item would arguably enhance the reliability of any ensuing identification or source attribution. In practice, however, obtaining handwriting “fillers” sufficiently similar to the questioned writing is far more difficult than obtaining suitable “fillers” for photo lineups, and such exemplar lineups may be of little usefulness in instances where the questioned writing displays several highly distinctive, individualizing writing features. In addition, double-blind lineups would seem to be far more feasible in handwriting investigations undertaken by public sector or institutional forensic laboratories where examiners can work with case managers possessing discipline-specific (handwriting) expertise and “blind” evidence lineup administrators than by private sector examiners who work as solo practitioners and receive their casework assignments directly from lawyers or clients, oftentimes accompanied by cues that indicate the desired outcome.

The use of single-blind exemplar lineups in handwriting cases is not a new development. For example, in investigating the source of anonymous handwritten letters emanating from a limited population of possible writers, experienced forensic document examiners routinely insist on using exemplar lineups and being kept blinded to which of the exemplars belongs to the suspected writer until such time as the examiner has completed all examinations and reached a decision.⁶³ However, the usefulness of such

⁶² See Canter, *supra* note 55, at 83–88.

⁶³ The failure to use exemplar lineups may also have contributed to the erroneous handwriting opinions offered in nineteenth-century France’s infamous Dreyfus Affair. Nearly a century later, France is once again confronted with the exemplar lineup issue in a murder case known as the “Gregory Affair.” The case involves the October 1984 kidnapping of four-year-old Gregory Villemin and a series of anonymous handwritten poison-pen letters that were sent to his family. After the boy’s body was discovered, a witness incriminated an uncle, who was indicted after a handwriting expert identified him as the author of the anonymous notes. The witness later recanted, but not before Gregory’s father killed the uncle to avenge his son’s murder. In a subsequent blind exemplar lineup procedure in which handwriting exemplars from all members of the family were examined and compared to the anonymous notes by a second expert, Gregory’s mother, Christine, was identified as the

lineups varies inversely with the distinctiveness of the handwriting features observed in the questioned writing(s); that is, the more distinctive the writing features in the questioned item, the more difficult it will be to find similar-looking “fillers,” and hence, the less useful the “fillers” will be.

Single-blind exemplar lineups prevent the handwriting expert from knowing the favored outcome because even if the expert has been exposed to other domain-irrelevant information, the expert would still not know which of the exemplars came from the suspected writer. Throughout my career, I have always preferred and made every reasonable effort to use single-blind exemplars in source attribution investigations, and I have publicly advocated for the routine use of single-blind exemplars in all source attribution handwriting investigations.⁶⁴

author, and she was indicted. Christine served almost eight years in prison for Gregory's murder before being cleared on appeal. See Richard Bernstein, *All of France is Asking: Who Killed Petit Gregory?*, N.Y. TIMES, (July 16, 1985), <http://www.nytimes.com/1985/07/16/world/all-of-france-is-asking-who-killed-petit-gregory.html>; see also INT'L HERALD TRIB., July 18, 1985, at 2, col. 3. In 2008, an appellate court ordered the case to be reopened in the hope that new advances in forensic science could shed light on DNA evidence. France still awaits the final chapter of this unsolved murder mystery that has seen two defendants indicted for the same murder on the basis of being identified by two different handwriting experts as the sole author of the anonymous notes.

⁶⁴ Andrew Sulner, *Why Forensic Scientists Should Embrace the Concept of Bias Control: A Practitioner's Perspective*, Speech at 67th Annual Meeting of the AAFS (Feb. 18, 2015) (video of speech viewable at <https://www.youtube.com/watch?v=QzH2tNVdqfw&feature=youtu.be>); see also Sulner, *Handwriting: Cognitive Bias*, *supra* note 36.

2018]

A DOCUMENT EXAMINER TELLS ALL

671

V. CASEWORK PEER REVIEWS: THE NEED TO KNOW WHAT THEY
ARE AND AREN'T AND WHY DIRECT TESTIMONY ABOUT ANY
REVIEW PERFORMED BY A NON-TESTIFYING EXPERT IS
INADMISSIBLE AS A MATTER OF LAW

A. *Types of Casework Peer Reviews*

There are three types of casework peer reviews that can be performed: (1) administrative reviews, (2) technical reviews, and (3) verifications.⁶⁵

1. Administrative Review

An *administrative review* involves a procedure used to check case file documentation and case reports for consistency with laboratory policy and for editorial correctness.⁶⁶ Essentially, it checks that the final report is coherent, reflects the examinations performed, and is free of spelling and grammar errors. It is acceptable for administrative reviews to be undertaken by someone without expertise in the relevant subject matter, i.e., forensic document examination.

2. Technical Review

A *technical review* involves a review of notes, data and other supporting documents that form the basis for a scientific conclusion to ensure that proper protocols and standards have been followed and that the case (bench) notes and any diagrams support the conclusion reported.⁶⁷ This type of review does not involve a full reexamination of the actual evidence; it simply checks that the correct and appropriate procedures have been followed and documented by the examiner, that the conclusions reached by the examiner are supported by the observations/results documented in the case file, and that any limitations with respect to the findings have been recorded and included in the report. Technical reviews must therefore be carried out by someone who is qualified in the relevant discipline, i.e., another FDE.⁶⁸

⁶⁵ Expert Working Group on Human Factors in Latent Print Analysis, *Latent Print Examination and Human Factors: Improving the Practice Through a Systems Approach*, NAT'L INST. OF JUST. & NAT'L INST. OF STANDARDS & TECH. 184 (Feb. 2012), http://ws680.nist.gov/publication/get_pdf.cfm?pub_id=910745 [hereinafter 2012 NIST LATENT PRINT REPORT]. See also AM. SOC'Y OF CRIME LAB. DIRS. & LAB. ACCREDITATION BD., ASCLD/LAB-INTERNATIONAL SUPPLEMENTAL REQUIREMENTS FOR THE ACCREDITATION OF FORENSIC SCIENCE TESTING LABORATORIES (2011), http://des.wa.gov/sites/default/files/public/documents/About/1063/RFP/Add7_Item4ASCLD.pdf; Scientific Working Grp. on Friction Ridge Analysis, Study & Tech., *Quality Assurance Guidelines for Latent Print Examiners: Version 4.0*, (Sept. 2006), http://clpex.com/swgfast/documents/quality-assurance/100210_Quality-Assurance_4.0.pdf.

⁶⁶ 2012 NIST LATENT PRINT REPORT, *supra* note 65, at 184.

⁶⁷ *Id.*

⁶⁸ *Id.* However, laboratory accreditation standards may not require that the individual

A *blind technical review* is one in which the reviewer has no knowledge of the examiner's conclusion prior to forming his own opinion, which is based solely on the case file documentation;⁶⁹ however, the reviewer may or may not have been blinded to domain-irrelevant case information.

3. Verification

This is a review procedure in which a second examiner independently examines all the evidence examined by the first examiner and reaches his or her own conclusion. A *blind verification* is a full-fledged reexamination of the same evidence by another expert who is blinded from domain-irrelevant context information and is unaware of the original examiner's conclusion.⁷⁰ A *double-blind*⁷¹ verification refers to a verification in which the reviewer does not know the identity of the initial examiner and *vice versa*.⁷² Government forensic handwriting laboratories frequently perform administrative and technical reviews as part of one review process, using a checklist or worksheet that identifies the elements of the review process. An example of an actual worksheet (redacted) used by a state crime laboratory appears on the next page:

conducting the technical review be an active forensic document examiner or currently be proficiency tested. See, e.g., AM. SOC'Y OF CRIME LAB. DIRS. & LAB. ACCREDITATION BD., *supra* note 65.

⁶⁹ 2012 NIST LATENT PRINT REPORT, *supra* note 65, at 105.

⁷⁰ See Canter, *supra* note 55, at 87; Inman, *supra* note 55, at 542.

⁷¹ "[A]n experimental procedure in which neither the subjects nor the experimenters know the make-up of the tests and control groups." WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY 74a (1993).

⁷² *Id.*

2018]

A DOCUMENT EXAMINER TELLS ALL

673

MARYLAND STATE POLICE - FORENSIC SCIENCES DIVISION	
TRACE ANALYSIS COVER SHEET	
Laboratory Number: _____	Inv. Agency Number: _____
Defendant: _____	Agency: _____
Date Received: <u>6/11/14</u>	Date(s) Analyzed: <u>6/12/14 - 7/9/14</u>
Analyst: _____	Officer: _____
Total Number of Pages in Case File: _____	Number of Items: _____
I have technically reviewed the following:	
<input checked="" type="checkbox"/>	Worksheet, notes, chain of custody
<input checked="" type="checkbox"/>	Appropriate examinations performed and procedures followed
<input checked="" type="checkbox"/>	Appropriate standards used for comparison
<input checked="" type="checkbox"/>	Documentation supports the findings
<input checked="" type="checkbox"/>	Clarity of notes in case of absence of Examiner
<input checked="" type="checkbox"/>	Where applicable, data or calculations were transcribed correctly
<input checked="" type="checkbox"/>	Associations are properly qualified in the test report
<input checked="" type="checkbox"/>	The test report contains all required information
Reviewer's signature: _____	Date: <u>7-9-14</u>
I have administratively reviewed the following:	
<input checked="" type="checkbox"/>	Checked for typos and grammar, description of items, analysis, results
<input checked="" type="checkbox"/>	The accuracy of unique identifying numbers and case numbers
<input checked="" type="checkbox"/>	The lab number and examiner's initials are recorded on every page.
<input checked="" type="checkbox"/>	Pages are numbered appropriately.
<input checked="" type="checkbox"/>	All notes, charts, photos, negatives, communication records, and description of evidence packaging and seals included
<input checked="" type="checkbox"/>	Evidence inventory and Chain of Custody
<input checked="" type="checkbox"/>	Documentation of verification for all positive associations.
<input checked="" type="checkbox"/>	If not checked in Technical Review, data or calculations were transcribed correctly.
<input checked="" type="checkbox"/>	The test report contains all required information
Reviewer's signature: _____	Date: <u>7/9/14</u>
Approving Authorities: Director, Deputy Director, Trace Evidence Supervisor, QA/Safety Manager	
Issue date: 3/7/2011	Page ____ of ____

B. *To Verify or Not to Verify?—Should the “V” in “ACE-V” Be a Prerequisite to the Admissibility of Handwriting Identification Opinion Evidence?*

“ACE-V” is the acronym used to describe a four-step process used in the forensic analysis of handwriting and other types of pattern evidence, most notably fingerprints; it stands for Analyze, Compare, Evaluate, and Verify.⁷³

“‘[M]any sources have described the verification phase of the ACE-V process to be a repeat of the ACE process done by another examiner,’ while ‘[o]ther sources describe verification as a confirmation of the original examiner’s conclusion.’”⁷⁴ Essentially, the “V” in “ACE-V” is the verification type of casework peer review defined above—a full and complete reexamination of the evidence by a second examiner that results in an independent verification of the first examiner’s reported findings and opinion.

Although the independent verification step of the ACE-V methodology is not a formal requirement for handwriting analysis, there are differing views as to whether it should be a prerequisite to the admissibility of handwriting identification opinion evidence. As both a practitioner of the discipline and a lawyer, my views on this subject necessarily consider both the practical and legal implications of mandating verification.

Most recently, this issue came to the forefront in two federal court cases, *Almeciga v. Center for Investigative Reporting*⁷⁵ and *Crew Tile Distrib., Inc. v. Porcelanosa Los Angeles, Inc.*⁷⁶ Both cases involved the same putative handwriting expert, Wendy Carlson, and the same task, evaluating the authenticity of a contested signature.⁷⁷ In both cases, motions were filed to

⁷³ See *United States v. Baines*, 573 F.3d 979, 983 (10th Cir. 2009).

⁷⁴ *Crew Tile Distribution, Inc. v. Porcelanosa L.A., Inc.*, No. 13-cv-3206, 2016 WL 8608447, at *4 (D. Colo. Sept. 12, 2016) (quoting Michele Triplett & Lauren Cooney, *The Etiology of ACE-V and Its Proper Use: An Exploration of the Relationship Between ACE-V and the Scientific Method of Hypothesis Testing*, 56 J. FORENSIC IDENTIFICATION 345 (2006)).

⁷⁵ *Almeciga v. Ctr. for Investigative Reporting, Inc.*, 185 F. Supp. 3d 401, 409 (S.D.N.Y. 2016).

⁷⁶ *Crew Tile Distribution*, 2016 WL 8608447.

⁷⁷ Ms. Carlson is not board certified by either the ABFDE or the BFDE, but her Curriculum Vitae included in an August 16, 2016 Expert Report she submitted in a New Jersey case (on file with the author) lists her as a “Certified Forensic Document Examiner and Registered Investigator,” and claims that “[i]n the last seven years, [she] has examined more than 10,000 documents and rendered opinions in approximately 1,000 active cases and multiple peer reviews involving questioned signatures, altered documents, handwritings, legal contracts, court documents, anonymous writing, and graffiti.” As noted in *Almeciga* and *Crew Tile*, her apprenticeship training took the form of a two-year part-time internet course which involved about five to ten hours of work per week under the tutelage of a mentor she met with personally when they were “able to connect.” The course was offered through the International School of Forensic Document Examination, an unaccredited school run by Bart Baggett, possibly with involvement by his father, Curtis Leo Baggett, that offers a “certification” upon completion of the course. Carlson’s CV also claims she obtained

exclude Ms. Carlson's testimony and opinions in their entirety, alleging, *inter alia*, that Carlson's failure to perform the verification step of ACE-V rendered her methodology critically flawed and her opinion consequently unreliable and inadmissible under Federal Rule of Evidence 702.

In *Almeciga*, U.S. District Court Judge Jed S. Rakoff of the Southern District of New York excluded Carlson's testimony, finding her "methodology [was] fundamentally unreliable and critically flawed," in part, because she failed to perform the verification step of the ACE-V process that she claimed to have followed. Judge Rakoff stated,

while Carlson purported to apply the ACE-V method in her expert report, . . . she admitted at the *Daubert* hearing that she did not have time to obtain a verification of her opinion in this case and that her report was inaccurate in this respect. Virtually by definition, then, Carlson failed to "reliably appl[y] the principles and methods" in question "to the facts of this case."⁷⁸

Judge Rakoff did not elaborate on why or whether he considers the lack of verification, in and of itself, to be such a critical flaw that it, alone, can be the death knell to admissibility, although some may take his decision to imply just that. Given the case-specific facts of *Almeciga*, and the counter-intuitiveness of assigning such critical importance to independent verification of an expert's opinion, I am inclined to believe that the lack of verification mentioned by Judge Rakoff was meant to highlight one of several striking contradictions between Carlson's reports and her in-court testimony, rather than to declare verification a *sine qua non* to a reliable handwriting identification opinion.

In *Crew Tile Distribution*, decided four months after *Almeciga*, U.S. District Court Judge William J. Martínez of the District of Colorado denied Plaintiff's Federal Rule of Evidence 702 motion to exclude Carlson's testimony.⁷⁹ In rejecting Plaintiff's argument that Carlson's failure to complete the verification step of ACE-V renders her opinions unreliable and inadmissible, Judge Martinez concluded that although the lack of verification is concerning, it does not render Carlson's testimony

"certifications" from the American College of Forensic Examiners for completion of their Registered Investigator course, Crime Scene Investigation course, and Digital Forensics Introduction course. For information on the American College of Forensic Examiners and the demise of its founder, see Radley Balko, *The Emperor of Junk Science Forensics Has Died*, WASH. POST (Aug. 31, 2017), https://www.washingtonpost.com/news/the-watch/wp/2017/08/31/the-emperor-of-junk-science-forensics-has-died/?utm_term=.402e4d5c0e69.

⁷⁸ Judge Rakoff did, however, note that "[i]n her Supplemental Expert Report, Carlson entirely drops the 'verification' step from her methodology, and purports to apply the 'ACE' methodology." *Almeciga*, 185 F. Supp. 3d at 426 n.14 (citations omitted).

⁷⁹ *Crew Tile Distribution*, 2016 WL 8608447, at *13.

inadmissible. The following excerpts reflect a well-reasoned judicial perspective with respect to the importance of verification in evaluating the reliability of handwriting identification opinions:

Plaintiff contends that Ms. Carlson's failure to complete the verification step renders her opinions unreliable and inadmissible. Plaintiff relies on *United States v. McDaniels*. There, the court concluded that the proponent of the evidence had not provided "clear evidence that [the expert] actually complied with steps 2–4 of this [ACE–V] methodology in the preparation of her expert report and testimony," and therefore struck the testimony for lack of a sufficient showing of reliability. *McDaniels* is not on all fours with the facts here, where Ms. Carlson completed the "ACE" steps, but not verification.

In performing the Court's gatekeeping function, "testing" and "peer review" are only two of the illustrative factors articulated in *Daubert*, and "independent testing is not the sine qua non of admissibility under *Daubert*." Moreover, "some of the *Daubert* factors may be less helpful when the evidence under consideration is not scientific in the strict sense." The gatekeeping function should be flexible, and "the relevant reliability concerns may focus upon personal knowledge or experience." Further, the Tenth Circuit, has criticized the verification step of ACE-V analysis as "not truly independent," suggesting that such verification adds little to reliability, but nevertheless finding expert testimony based on ACE-V protocols admissible.

Given these considerations, while the lack of verification is concerning, the Court concludes that it bears on the weight and credibility of Ms. Carlson's testimony. It does not render the testimony inadmissible. The Court's role as gatekeeper is not to determine whether the opinion is correct, only that the method was reliable enough to be presented to the jury. As described in the materials submitted by Plaintiff, the role of verification in this context is primarily to make the expert's work "reviewable," even if it was not actually reviewed. Plaintiff does not argue that Ms. Carlson's work is not "reviewable," nor that an expert of Plaintiff's choosing could not have reviewed and critiqued the analysis based on the work shown in her report and supporting materials.

Moreover, as a practical matter, making the verification step of ACE-V a prerequisite in every case would essentially require parties to obtain two experts' opinions before the first became

admissible. And while verification might cause an examiner to revisit her approach, it does not directly alter the examination and review completed under the first three steps of ACE-V. The Court is therefore not persuaded that obtaining verification was a prerequisite to admissibility in this case. The Court therefore agrees with the Fourth Circuit's conclusion in *Crisp*, that to the extent Ms. Carlson's analysis is "flawed or flimsy," including for lack of verification, "an able . . . lawyer will bring that fact to the jury's attention," and the jury will be "left to examine the [Agreement] and decide for itself whether it agree[s] with the expert."⁸⁰

The different perspectives of Judges Rakoff and Martinez probably have more to do with the specific facts that came to light in the respective *Daubert* hearings, rather than a fundamental difference of opinion as to how Rule 702 should be applied. However, this is an issue that requires an in-depth analysis of the respective facts of each case, something that is beyond the scope of this Article.⁸¹ In any event, I believe this issue is somewhat of a "red herring" and moot because it is now settled law that verification of a testifying expert's opinion by a non-testifying expert may not be brought up in the direct examination of the testifying expert,⁸² as discussed below.

C. *Testimony About Peer Review(s) Performed by a Non-Testifying Expert Constitutes Improper Bolstering and/or Inadmissible Hearsay*

FDEs, as well as expert witnesses from other forensic science disciplines, have often testified that their casework is subjected to so-called "peer review." The proponent of such testimony seeks its admission to demonstrate the quality assurance protocol employed in the testifying expert's forensic laboratory prior to the issuance of a final report.⁸³ However, implicit in any such proffer is the notion that the peer review process enhances the reliability of the proffered opinion evidence. The typical line of inquiry posed to the testifying expert might be:

⁸⁰ *Crew Tile Distribution*, 2016 WL 8608447, at *4–5 (citations omitted).

⁸¹ Nevertheless, in Part VI of this Article, *infra*, I have provided some limited, case-specific analysis with respect to a critical issue in that case—the impropriety and unreliability of Ms. Carlson's exclusive reliance on self-serving *post litem motem* exemplars in concluding that the contested signature was a forgery, especially in the face of facts strongly suggesting that the disclaiming signatory had provided Ms. Carlson with disguised exemplars.

⁸² However, the testifying expert may not be precluded from mentioning it in response to a question posed on cross-examination that asks if his or her work was subjected to any kind of "peer" review by another examiner.

⁸³ Casework reviews are designed to address quality of output and human factors issues, which can result in nonconformity with laboratory procedures and documentation requirements, and errors in report writing and testimony.

Q: Was your work peer reviewed by anyone else?

A: Yes. Our lab has 100% peer review, both technically and administratively, by colleagues within the same section. Every case worked on by one examiner is reviewed by another, and if there is any disagreement between the two, a report does not leave the lab.

Lawyers and trial judges mistakenly assume that the mere mention of the words “peer review” equates to a comprehensive reexamination of the evidence and an independent verification of a given opinion or conclusion. This is because they are unaware of the different types of reviews and which of them is or is not routinely performed.⁸⁴ Unless one knows what a technical or administrative review really entails, the above answer is misleading and disingenuous. The above answer clearly implies that before finalizing any report or conclusion, a “second set of eyes” reviewed the very same evidence and came to the same conclusion as the testifying expert, which can only occur with a verification type of review.

Trial judges have routinely allowed this type of testimony and appellate courts have typically found that any error in doing so was harmless⁸⁵—until recently.⁸⁶ Some lawyers have attempted to challenge such testimony by

⁸⁴ A technical review is sometimes referred to as peer review in government forensic laboratory settings. The term “peer review” is more widely used to describe the process of reviewing manuscripts submitted for publication to a scientific journal, which is customarily double-blind.

⁸⁵ *Bunche v. State*, 5 So. 3d 38, 41 (Fla. Dist. Ct. App. 2009). *See id.* (determining error was harmless when fingerprint handwriting expert testified that a non-testifying expert verified his work); *State v. Jones*, 368 S.E.2d 844, 846–49 (N.C. 1988) (finding trial court properly permitted a fingerprint expert to testify that another expert had checked and concurred with the testifying expert’s conclusion, because under the standard procedures followed by the expert he could not have arrived at and testified to his opinion without the verification by the other expert); *State v. Williams*, No. 95CA93, 1996 WL 753216, at *10 (Ohio Ct. App. Dec. 3, 1996) (permitting fingerprint expert’s testimony that he had his results verified by another fingerprint expert).

⁸⁶ *See Miller v. State*, 127 So. 3d 580 (Fla. Dist. Ct. App. 2012) (determining error was harmful when handwriting expert testified that non-testifying expert also determined handwriting belonged to defendant). *See also Potts v. State*, 57 So. 3d 292, 294 (Fla. Dist. Ct. App. 2011) (determining error was harmful when fingerprint analyst testified that another fingerprint expert determined fingerprints belonged to defendant); *Telfort v. State*, 978 So. 2d 225, 227 (Fla. Dist. Ct. App. 2008) (holding that the trial court committed reversible error by allowing a fingerprint expert to testify that his identification of the defendant’s fingerprint was verified by two other examiners); *State v. Smith*, 628 N.E.2d 1176, 1181 (Ill. App. Ct. 1994) (finding fingerprint technician’s testimony that her identification was verified by another technician was hearsay); *State v. Connor*, 937 A.2d 928, 930–32 (N.H. 2007) (finding fingerprint expert’s testimony regarding another expert’s verification inadmissible, because the verification did not form a basis for the testifying expert’s opinion, but was simply a necessary prerequisite to the release of his already formed opinion); *State v. Wicker*, 832 P.2d

establishing that the casework peer review was biased and unreliable because it was not performed blindly, e.g., the reviewer was a “friendly” colleague or coworker, working in the very same laboratory unit or office. In the past, I have publicly advocated that lawyers should instead seek to exclude such testimony altogether because it constitutes improper bolstering or inadmissible hearsay, or both,⁸⁷ an argument that has found judicial acceptance of late.⁸⁸ Accordingly, lawyers, judges, and testifying experts need to know the impact of recent court rulings declaring that testimony about case review performed by non-testifying experts is inadmissible on evidentiary grounds.

Before addressing the cases, I must first confess that my previous approach (and that of many lawyers and courts) was too narrow in that it failed to sufficiently emphasize the distinction between bolstering and hearsay. “Bolstering” refers to the proffer of information for the sole purpose of raising the credit the jury might give to a witness’s evidence before the witness’s credibility has been attacked in any way on cross examination or otherwise.⁸⁹ The most common example of bolstering evidence is evidence of a prior consistent statement to rebut a charge of recent fabrication.⁹⁰ Such evidence is not admissible unless and until there is some suggestion by the opponent on cross-examination or otherwise that a witness’s testimony on direct is a recent fabrication.⁹¹ Then evidence of prior consistent statements by the witness that occurred before the suggested impetus to falsify are admissible to rehabilitate the witness. Traditionally, there was no hearsay exception for such prior consistent statements, but they

127, 128–30 (Wash. Ct. App. 1992) (finding fingerprint expert’s testimony that his identification was verified by another technician was inadmissible hearsay).

⁸⁷ Andrew Sulner, *Impeaching Self-Serving Case Review Evidence and Excluding Case Review Testimony by a Second Forensic Document Examiner*, Speech at the Annual Symposium of the Association of Forensic Document Examiners (Nov. 6, 2009); Andrew Sulner, *Disingenuous Testimony About Casework Peer Review: Understanding What it Really Means and When Such Testimony Will Be Precluded*, Speech at the 64th Annual Meeting of the American Academy of Forensic Sciences (Feb. 21, 2012); Andrew Sulner, *Legal Update: Recent Court Decisions Declaring Testimony About Case Reviews Performed by Non-Testifying Experts to be Inadmissible*, Speech at the 67th Annual Meeting of the American Academy of Forensic Sciences (Feb. 20, 2015); Andrew Sulner, *Peer Reviews of Forensic Casework: Understanding the Different Types of Standard Reviews and Why Courts Have Ruled that Evidence of Casework Peer Reviews Performed by Non-Testifying Experts is Inadmissible*, Speech at the Annual Symposium of the Association of Forensic Document Examiners (Oct. 29, 2016).

⁸⁸ See *supra* note 86.

⁸⁹ See generally *IMWINKELRIED ET AL.*, *supra* note 3, § 702.

⁹⁰ There may be other circumstances where prior consistent statements are considered allowable rehabilitation, see FED. R. EVID. 801(D)(1)(b)(ii), but the illustration is confined to the rebuttal of recent fabrication as set out by FED. R. EVID. 801(D)(1)(b)(i).

⁹¹ See *IMWINKELRIED ET AL.*, *supra* note 3, § 720. Professor Imwinkelried notes that some jurisdictions also allow rehabilitation by prior consistent statement when there is a suggestion that the live witness’s memory is faulty on direct examination.

were admitted on the assertedly non-hearsay ground that independent of the truth of either the statement from the witness stand or the prior statement, the correspondence shows that the assertions were not made up as a result of the suggested impetus to falsify.⁹² All purely accrediting information in the form of an assertion by anyone, whether the witness or another, is admitted only on such a non-hearsay rationale if there is no hearsay exception or exclusion covering it.⁹³ When the information only accredits the witness if it is true, this is merely corroboration, and the proper objection is not “improper bolstering,” but “improper corroboration by hearsay assertions not within any exception.”⁹⁴ It is hard to formulate a credible non-hearsay rationale for the admission of opinions by other experts that the testifying expert’s procedures were proper, or that the testifying expert’s opinions were accurate. So the proper objection should probably be “improper bolstering and improper corroboration by hearsay not within any exception,” supplemented in a criminal case by an objection to the violation of the Confrontation rights of the defendant⁹⁵ if the witness is a witness for the prosecution.

In *Miller v. State of Florida*,⁹⁶ the District Court of Appeal for the Fourth Circuit of Florida overturned a defendant’s conviction for robbery and murder based, in part, on the trial court having permitted two government handwriting experts to testify that their conclusions were subject to a peer review process in which another analyst had corroborated their findings after independently examining and analyzing the same evidence. The Florida appellate court’s decision was based upon the evidentiary rules prohibiting testimony bolstering and the constitutional right of Confrontation guaranteed by the Sixth Amendment.⁹⁷

The *Miller* case involved a restaurant supervisor who was found duct taped to a folding chair in the closed restaurant and shot in the head.⁹⁸ The only physical evidence linking the disgruntled ex-employee, Jesse Miller, to the crime scene was a small amount of handwriting on a note pad that

⁹² *Id.*

⁹³ Because instructions trying to make this distinction to the jury are ineffective in regard to prior consistent statements of the witness him- or herself, most modern evidence codes provide for a hearsay exception for qualifying prior consistent statements (or an exclusion from being classified as hearsay, which is functionally the same thing). See FED. R. EVID. 801(D)(1)(b) (stating a hearsay exclusion); N.J. R. EVID. 803(a) (stating a hearsay exception).

⁹⁴ I thank Edward Imwinkelried for pointing this out to me.

⁹⁵ It is hard to imagine how such assertions might be categorized as “non-testimonial” as that phrase is used in *Crawford v. Washington* and its progeny so as to escape confrontation requirements.

⁹⁶ *Miller v. State*, 127 So. 3d 580 (Fla. Dist. Ct. App. 2012).

⁹⁷ *Id.* at 586.

⁹⁸ *Id.* at 581.

2018]

A DOCUMENT EXAMINER TELLS ALL

681

contained the following writing: “‘55, 65, 9, 10,’ ‘4 time stop,’ ‘left right left right.’”⁹⁹

The issue on appeal was whether the trial court had erred in allowing two handwriting experts to bolster their testimony by describing the peer review process.¹⁰⁰ The trial judge had overruled defense counsel’s objection on the ground of “bolstering” and allowed the first handwriting expert (FDE #1) to give the following testimony:

Q. Okay. Does your conclusion have to be peer reviewed by another analyst within the context of your laboratory?

A. It does.

Q. And is that part of your protocols and procedures?

A. Yes.¹⁰¹

FDE #1 concluded with a “high degree of probability” that the defendant executed the text on the yellow notepad, explaining that “high degree of probability” meant it was a virtual certainty that the defendant had written the text. The State then asked if his conclusion was contained in the report issued by the lab:

Q. Now, let me ask you this question: Was that the opinion within the purview of your laboratory protocols?

A. The ultimate conclusion that I just gave, the high degree of probability, was the conclusion that exited the laboratory due to the peer review process that we follow.

Defense Counsel: Objection, bolstering.

Court: Overruled.¹⁰²

⁹⁹ *Id.* at 581. “A partial print, sufficient for identification, was obtained from the duct tape found on the victim. No identification was made despite numerous comparisons, including the defendant’s. The latent fingerprints obtained from various surfaces at the crime scene matched those of other employees, but not the defendant. The only latent fingerprint on the yellow notepad belonged to the restaurant manager.” *Id.* at 581–82.

¹⁰⁰ *Id.* at 585.

¹⁰¹ *Id.* at 583.

¹⁰² *Miller*, 127 So. 3d at 583.

FDE #1 then went on to clarify that his initial conclusion was a “positive” identification, which was a higher level of certainty than that contained in the final report, and that it was only through the peer review process that the initial conclusion went down one level to “high degree of probability.”¹⁰³

The State then presented the testimony of a second examiner (FDE #2), who testified that in his expert opinion, the defendant wrote the words and numbers on the yellow notepad to the exclusion of everyone else.¹⁰⁴ He made a full identification, the highest conclusion in the field of forensic document examination.¹⁰⁵ The State then had FDE #2 testify about the peer review of his findings, once again over defendant’s objection that such testimony constituted improper bolstering, as reflected by the following exchange:

Q. When you formulate that opinion within the context of the laboratory, for whom you worked at the time that you did this work, are there quality assurance controls in place within the context of that laboratory?

A. Yes, we have—

Defendant: Objection, Judge, bolstering.

Court: Okay, I’m sorry. The objection is bolstering?

Defendant: Bolstering, I’m sorry.

Court: Overruled.

Q. And could you explain to the members of the jury what those quality assurance controls are?

A. We have two different types of quality assurance processes, one is a technical review and one is a confirmation. Technical review, every tenth case normally that is examined is reexamined by another examiner to ensure that the protocols have been followed, the standards have been applied, and that the notes and findings or that the notes, diagrams or whatever actually support the conclusion that is being offered. Additionally, in our laboratory, whenever there was an identification made, a separate examiner is required to independently examine the evidence associated with that case and reach his or her own conclusion. The

¹⁰³ *Id.*

¹⁰⁴ *Id.* at 584.

¹⁰⁵ *Id.*

2018]

A DOCUMENT EXAMINER TELLS ALL

683

identification would only go out of the laboratory based on our own protocols if there's confirmation.

Q. And did you have that in this case?

A. Yes, sir, I did.¹⁰⁶

On appeal, the defendant argued that the trial court abused its discretion by allowing the State's two document examiners to testify about peer reviews performed by other non-testifying examiners.¹⁰⁷ In reversing the defendant's conviction, the appellate court agreed with the defense that the peer review testimony constituted improper bolstering and was not harmless, providing the following elucidating legal analysis:

In *Schwarz v. State*, we held that an expert may not testify that he formed his opinion by consulting with others in the same field. Subsequently, we extended the reach of that concept to preclude an expert from bolstering—his own opinion on direct examination with that of another expert.” We explained that this prohibitive rule applies equally whether the expert relies on a second expert's opinion or testifies about other experts' opinions in explaining the process employed.

Here, the defendant specifically lodged a bolstering objection to both examiners testifying about the process and called the court's attention to our decision in *Bunche*. Nevertheless, the court overruled the objections and permitted the testimony. In doing so, the trial court erred. We must next decide if the error was harmless.

In *State v. DiGuilio*, our supreme court explained that: The harmless error test . . . places the burden on the state . . . to prove beyond a reasonable doubt that the error complained of did not contribute to the verdict or . . . that there is no reasonable possibility that the error contributed to the conviction.

Here, as in *Bunche*, the handwriting examiners testified that they “use[d] . . . a second examiner in the verification process.” The first examiner testified that, “in our laboratory, whenever there was an identification made, a separate examiner is required to independently examine the evidence associated with that case and

¹⁰⁶ *Id.*

¹⁰⁷ *Id.* at 585.

reach his or her own conclusion. *The identification would only go out of the laboratory based on our protocols if there's confirmation.*" The second examiner testified that, "the high degree of probability . . . was the conclusion that exited the laboratory due to the peer review process that we follow." Both of these comments improperly bolstered the testifying experts' findings because they relied on the "use of a second examiner in the verification process."

The State argues that no improper bolstering occurred because the experts "were simply providing a general explanation of the [peer review] process." While that may be true, it does not eliminate the harm of admitting the opinions of non-testifying experts to bolster the testimony of those testifying. Instead, it deprives the opposing party of the opportunity to cross-examine the non-testifying experts. This was the reasoning in *Bunche*, notwithstanding that we found the error harmless there.

The State attempts to distinguish *Telfort* and *Bunche* because it presented two independent experts instead of one. The State argues the expert examiners' testimony was properly admitted because their conclusions "were consistent with each other." Two wrongs do not make a right. Two instances of inadmissible bolstering testimony do not ameliorate the harm and prejudice simply because the experts come to the same conclusion. Rather, two instances double the harm.

The State also suggests that *Telfort*, *Bunche*, and *Potts v. State* conflict with cases from the Third and Fifth Districts. We disagree. *J.V. and Masters* involved expert witnesses rendering an opinion after reviewing "facts and data." Such testimony is entirely permissible under the Florida Evidence Code. There simply is no conflict.

Because the defendant contested the experts' conclusions that he wrote the words and numbers on the yellow notepad, and those opinions are the only real direct evidence placing the defendant at the scene of the crime, we cannot find the error harmless.¹⁰⁸

¹⁰⁸ *Id.* at 585–86 (citations omitted). In *Bunche*, the court held the error to be harmless because there was no dispute that the prints belonged to the defendant—the question was when the defendant's prints were left at the scene. *Bunche v. State*, 5 So. 3d 38, 41 (Fla. Dist. Ct. App. 2009).

However, despite the well-reasoned opinion of the Florida District Court of Appeal, prior case law from other states reflects differing opinions on whether testimony about verification by a non-testifying expert should be allowed or excluded on evidentiary grounds.¹⁰⁹

VI. GATEKEEPING: TRIAL JUDGES AND LAWYERS ARE STILL LARGELY INEFFECTIVE IN PREVENTING THE ADMISSIBILITY OF UNRELIABLE OR EXAGGERATED EXPERT TESTIMONY

The 2009 NAS Report recommended that the handling of forensic science evidence in court be vastly improved, emphasizing the need to challenge poor and marginal evidence from traditional forensic science disciplines and to prevent the admissibility of unreliable or exaggerated opinion evidence.¹¹⁰ Unfortunately, very little improvement has been seen in this area of great concern to the interests of justice.¹¹¹ This may be due to how members of the bench and bar have traditionally viewed their respective roles. Although all judges will acknowledge the need to exclude expert opinions that are so weak or speculative that they would tend to mislead or confuse the jury, many still seem to have difficulty in distinguishing between opinion evidence that is shaky but admissible from that which is unreliable and inadmissible.¹¹² Perhaps out of fear of being reversed on appeal, judges may be too reluctant to exclude proffered evidence as being unreliable, preferring instead to let that issue be resolved through the traditional means of attacking weak or marginal evidence: cross-examination and presentation of contrary proof.¹¹³ Regrettably, experience has shown that far too many

¹⁰⁹ Prior decisions from other states disallowing such testimony include: *Bell v. State*, 724 S.W.2d 780, 800–01 (Tex. Crim. App. 1986) (explaining the question of whether a fingerprint expert’s ID of a print was verified by another expert is improper since it is an attempt to bolster the testifying expert’s testimony, although such a question would be proper as a predicate to introducing the second technician’s analysis); *State v. Wicker*, 832 P.2d 127, 128–30 (Wash. Ct. App. 1992); *State v. Smith*, 628 N.E.2d 1176, 1181 (Ill. App. Ct. 1994); *State v. Connor*, 937 A.2d 928, 930–32 (N.H. 2007) (explaining fingerprint expert’s testimony regarding another expert’s verification is inadmissible hearsay, as the verification did not form a basis for the testifying expert’s opinion, but was simply a prerequisite to the release of his already formed opinion). Two prior decisions allowing such testimony are *State v. Jones*, 368 S.E.2d 844, 846–49 (N.C. 1988) (finding trial court properly permitted a fingerprint expert to testify that another expert had checked and concurred with the testifying expert’s conclusion, because under the standard procedures followed by the expert, he could not have arrived at and testified to his opinion without the verification by the other expert); *State v. Williams*, No. 95CA93, 1996 WL753216 (Ohio Ct. App. Dec. 3, 1996) (permitting fingerprint expert’s testimony that he had his results verified by another fingerprint expert).

¹¹⁰ See 2009 NAS REPORT, *supra* note 43, at 4, 48.

¹¹¹ Keelah E.G. Williams & Michael J. Saks, *Why Don’t the Gatekeepers Guard the Gates?: Comments Prompted by Edmond*, 36 ADELAIDE L. REV. 109–24 (2015).

¹¹² *Id.*

¹¹³ Even the Advisory Committee Note to FRE 702 provides that “rejection of expert testimony is the exception rather than the rule” and that “[t]he trial court’s role as gatekeeper is not intended to serve as a replacement for the adversary system.” FED. R. EVID. 702

trial lawyers are ill-equipped to challenge unreliable expert testimony, especially in the criminal justice system, where defense attorneys lack the resources needed to obtain the expertise required to successfully challenge such testimony.

A. *The Responsibility of Gatekeepers Regarding Proffered Expert Testimony*

Federal Rule of Evidence Rule 702 (FRE 702) provides:

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:

- a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
- b) the testimony is based on sufficient facts or data;
- c) the testimony is the product of reliable principles and methods; and
- d) the expert has reliably applied the principles and methods to the facts of the case.¹¹⁴

To be admissible under FRE 702, the expert and/or the lawyer offering the expert testimony must provide the trial judge with a warrant for believing that (a) the expert possesses sufficient skill with regard to performing the particular task at hand, and (b) reliable inferences can be drawn from the reliable performance of that particular task.¹¹⁵

As the gatekeeper entrusted with determining whether to admit or exclude expert testimony, scrutiny of the proffered expert's qualifications is but one factor that the trial judge must consider. Once the proffered witness is qualified as an expert, the trial court's gatekeeping function requires more than simply "taking the expert's word for it" and allowing the witness to testify simply because he or she was qualified as an expert. Before any

advisory committee's note to 2011 amendment.

¹¹⁴ FED. R. EVID. 702.

¹¹⁵ The Supreme Court's decision in *Kumho Tire Co. v. Carmichael* made it clear that the evaluation must be directed to the reliability of the expertise in the specific "task at hand." See *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 141 (1999).

proffered expert testimony can be admitted, the trial judge must also scrutinize the relevance and reliability of the proffered testimony.

Under FRE 702, as under *Daubert*¹¹⁶ and *Kumho*,¹¹⁷ trial courts must¹¹⁸ scrutinize not only the principles and methods used by the expert, but also whether those principles and methods have been properly applied to the facts of the case. When an expert purports to apply principles and methods in accordance with professional standards, yet reaches a conclusion that other experts in the field would not reach, the trial court may fairly suspect that the principles and methods have not been faithfully applied.¹¹⁹ “Any step that renders the analysis unreliable . . . renders the expert’s testimony inadmissible. This is true whether the step completely changes a reliable methodology or merely misapplies that methodology.”¹²⁰

B. Factors Relevant and Material to Properly Assessing the Reliability of Proffered Opinion Evidence Under Rule 702

The most effective way for a court to evaluate reliability and the likelihood of error associated with proffered expert testimony is to examine with critical common sense the methodology used and reasons given for a conclusion under the circumstances of the specific task at hand in the case.¹²¹ This requires evaluating the reasoning used in forming an expert opinion and determining if it is properly grounded, well-reasoned, and not speculative. As long as the court’s analysis focuses on the expert’s methods and reasoning, and not on the expert’s conclusions, its actions are proper.¹²²

Trial judges must consider various factors in determining whether expert testimony is sufficiently reliable to be considered by the trier of fact. The Advisory Committee’s Notes to Federal Rule of Evidence 702, as amended in 2011, describes three factors that are relevant and material to properly assessing the reliability of proffered opinion evidence as follows:

- a) whether the expert has unjustifiably extrapolated from an accepted premise to an unfounded conclusion;
- b) whether the expert has adequately accounted for obvious alternative explanations; and

¹¹⁶ *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993).

¹¹⁷ *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999).

¹¹⁸ Rule 702 not only vests the trial judge with the authority to scrutinize and evaluate the expert’s methodology and reasoning, it requires the trial judge to do so. *Claar v. Burlington N.R.R.*, 29 F.3d 499, 501 (9th Cir. 1994).

¹¹⁹ *See Lust v. Merrell Dow Pharm., Inc.*, 89 F.3d 594, 598 (9th Cir. 1996).

¹²⁰ *In re Paoli R.R. Yard PCB Litig.*, 35 F.3d 717, 745 (3d Cir. 1994).

¹²¹ This was the guidance offered by the Supreme Court in *Kumho*. *See Kumho*, 526 U.S. at 149.

¹²² *See Kennedy v. Collagen*, 161 F.3d 1226, 1230 (9th Cir. 1998).

- c) whether the expert “is being as careful as he would be in his regular professional work outside his paid litigation consulting.”¹²³

For trial judges to make a meaningful reliability assessment using the above criteria, they need to know the discipline’s best practices and guidelines for performing the task at hand, as set forth in consensus standards adopted by the profession, learned treatises, and relevant case law. Such knowledge is essential to performing the gatekeeping task of distinguishing between evidence that is shaky and admissible from that which is unreliable and inadmissible.

When an expert’s departure from an acknowledged professional standard or best practice can be shown to have produced an illogical or exaggerated opinion, the decision to exclude such an opinion should become relatively straightforward. However, a review of relevant appellate court decisions shows this to be easier said than done.

C. *Court Decisions Illustrating Unreliable and Critically Flawed Testimony on the Part of Presumptively Well-Qualified Handwriting Experts That Should Have Been Excluded by the Trial Judge*

1. Improper Methodology—Using Exemplar Signatures of a Remote Date to Support an Opinion of Forgery

Hardin v. Montgomery,¹²⁴ a Kentucky case involving allegations of election misconduct and violations of the Corrupt Practices Act, provides a good example of an appellate court conducting the necessary reliability assessment of handwriting identification opinion evidence after the trial court failed to do so. The trial court admitted expert testimony concerning voter signatures, with ABFDE Diplomate Thomas W. Vastrick¹²⁵ and BFDE

¹²³ FED. R. EVID. 702 advisory committee’s note to 2011 amendment (“*Daubert* requires the trial court to assure itself that the expert ‘employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.’” (quoting *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999))).

¹²⁴ *Hardin v. Montgomery*, No. 2015-CA-000305, 2015 Ky. App. LEXIS 94 (Ky. Ct. App. June 12, 2015).

¹²⁵ Mr. Vastrick was and is a presumptively well-qualified FDE, having been, *inter alia*, certified by the American Board of Forensic Document Examiners (ABFDE), Chair of the Questioned Document Section of the American Academy of Forensic Sciences (AAFS), and a longstanding member of the American Society of Questioned Document Examiners (ASQDE).

Diplomate Steven A. Slyter¹²⁶ testifying for opposite sides.¹²⁷ Mr. Vastrick testified that the signatures of forty-three voters did not match the signatures on their voter registration cards, which was the only reference exemplar he used for comparison purposes.¹²⁸ All the comparisons performed by Mr. Vastrick were based upon photocopies and consisted solely of one-to-one comparisons between the challenged voter signature and the signature appearing on the named individual's voter registration card. Of the thirty-eight dated signature pairs compared by Montgomery's expert (Mr. Vastrick), seven pairs had time differences of forty plus years, eleven pairs had time differences of thirty to thirty-nine years, four pairs had time differences of twenty to twenty-nine years, and seven pairs had time differences of ten to nineteen years.¹²⁹ FDE Slyter testified that the methodology employed by Mr. Vastrick violated established professional standards and best practices for performing handwriting comparisons and was wholly improper and unreliable.¹³⁰ Based in large part on Mr. Vastrick's testimony the trial court found, *inter alia*, voter fraud and other election law violations, and voided the election.¹³¹ The Kentucky Court of Appeals affirmed by a 2-1 vote¹³² over a blistering dissent by Judge Thompson, who, in evaluating Vastrick's methodology, found it to be fundamentally unreliable and critically flawed, stating:

¹²⁶ Mr. Slyter was and is a presumptively well-qualified FDE, having been, *inter alia*, certified by the Board of Forensic Document Examiners (BFDE), President of the BFDE, a longstanding member of the Association of Forensic Document Examiners (AFDE), and author of a well-recognized treatise on forensic signature analysis, *FORENSIC SIGNATURE EXAMINATION*, Charles C. Thomas Publisher, Springfield (1995).

¹²⁷ There are two boards that certify forensic document examiners, both of which are accredited by the Forensic Specialties Accreditation Board (FSAB), which was established in 2000 by the National Institute of Justice, the American Academy of Forensic Sciences, and the National Forensic Science Technology Center to serve the same function for forensic science disciplines as that served for medical specialties by the American Board of Medical Specialties. See *Forensic Specialties Accreditation Board, Inc.*, (Mar. 16, 2018), <http://thefsab.org/>. The Board of Forensic Document Examiners (BFDE) was the first board to be accredited by FSAB in 2006; the American Board of Forensic Document Examiners (AFDE) became FSAB-accredited in 2007. Document examiners certified by either of these Boards are designated as "Diplomates."

¹²⁸ *Id.* at *9.

¹²⁹ Forensic Document Examination Report of Thomas W. Vastrick, *Montgomery v. Magoffin Cty. Bd. of Elections*, No. 14-ci-00371, 2015 WL 4876464 (Ky. Cir. Ct. Feb. 20, 2015) (on file with author); Transcript of Record, *Montgomery*, 2015 WL 4876464 (Thomas W. Vastrick's Testimony) (on file with author).

¹³⁰ This was derived from court opinions (*see supra* note 124, *infra* note 134) and supplemented by information provided by Steven Slyter, including the Magoffin County Kentucky Circuit Court video of his testimony on February 9, 2015, which is on file with author.

¹³¹ *Hardin v. Montgomery*, No. 2015-CA-000305, 2015 Ky. App. LEXIS 94, at *36 (Ky. Ct. App. June 12, 2015).

¹³² *Id.* at 11.

Montgomery's handwriting expert, Thomas Vastrick, provided only marginally reliable evidence of forgeries. The other handwriting expert, Stephen Styler [sic], testified that Vastrick's one-to-one comparison of signatures was an unreliable method and could not establish any forgeries. Even our case law teaches that comparison of a single signature with a challenged signature is not a reliable method to determine the authenticity of the signature.¹³³

The Supreme Court of Kentucky granted certiorari to review the case, and on August 16, 2016, issued an opinion¹³⁴ reversing the Court of Appeals, and stating, *inter alia*:

Montgomery introduced at the trial the testimony of Thomas Vastrick, an expert on handwriting analysis who had examined voter signatures on the election day voter roster at the Flat Fork precinct and on the absentee ballot materials. Vastrick opined that the signatures of forty-three Flat Fork voters did not match the corresponding signature on the voter's voter registration card; that fourteen voter signatures on voting precinct forms did not match, and that twenty-six voter signatures on the absentee ballots did not match the corresponding signature on the voter's absentee ballot application. The implication of his opinion is that eighty-three votes cast in the names of those voters were cast by imposters who forged the signatures of the registered voter. Only two of the eighty-three voters were called as witnesses and both refuted the insinuated forgery. Both verified the authenticity of their signatures on the voting roster and attested to having personally cast the votes recorded in their names. One of the voters explained that her current signature might look different than the signature on her voter registration card because the latter was signed thirty years prior, when she was eighteen, and the former was signed on election day. The other voter testified that he was left-handed, and because he had a broken left arm on election day he had to sign the voter roster with his right hand. None of the other voters whose election day signature was identified by Vastrick as suspect were called to testify. Montgomery's attempt to demonstrate that imposters cast ballots in place of legitimate registered voters by forging their signatures falls woefully short. Proving the suspected forgeries would have been relatively easy because the names and addresses of the eighty-three voters whose signatures

¹³³ *Id.* at 14 (citation omitted).

¹³⁴ *Hardin v. Montgomery*, 495 S.W.3d 686 (Ky. 2016).

were suspect were readily available.

Appellants refuted Vastrick's opinion with the countervailing analysis by their handwriting expert, Stephen Styler [sic]. Ultimately, Vastrick conceded that he used an unreliable method of handwriting comparison and could not definitively establish any forgeries. In this vein, our predecessor court has acknowledged that the comparison of a single signature with a challenged signature is not a reliable method to determine the authenticity of the signature in question.¹³⁵

Unfortunately, Mr. Vastrick is not the only presumptively well-qualified handwriting expert to commit this fundamental type of error, and *Hardin* is not an isolated case of unreliable handwriting identification opinion evidence escaping detection and exclusion at the trial level. For instance, only a few years earlier, in another civil case involving questioned voter signatures, *Felder v. Storobin*,¹³⁶ a New York appellate court discredited the same type of unreliable handwriting testimony on the part of another presumptively well-qualified document examiner,¹³⁷ who likewise based an opinion of forgery on perceived differences appearing in signatures of a remote date.

2. Illogical (Unfounded) Conclusion—Unjustifiably
Extrapolating from an Accepted Premise to an Unfounded
Conclusion

*Adams v. Weber*¹³⁸ involved an action brought by Samuel D. Adams, a/k/a Dale S. White, pursuant to an application for a writ of habeas corpus. Adams claimed that his court-appointed defense attorney (Brankin) provided ineffective assistance of counsel when representing him in connection with a 2001 criminal case.¹³⁹

As described in the South Dakota Circuit Court's memorandum decision, the underlying criminal prosecution arose out of an incident that occurred while Gayle Wanous ("Ms. Wanous") was working alone in her

¹³⁵ *Id.* at 706–07 (citing *Beauchamp v. Willis*, 189 S.W.2d 938, 941 (Ky. 1945)).

¹³⁶ 953 N.Y.S.2d 602 (App. Div. 2012).

¹³⁷ The handwriting expert who testified in this case was employed for thirty years with a state police crime laboratory that allowed him to accept private sector civil casework; he was board-certified by the American Board of Forensic Document Examiners (ABFDE) and was a long-standing member of the Questioned Document Section of the American Academy of Forensic Sciences (AAFS) and the American Society of Questioned Document Examiners (ASQDE). The expert's curriculum vitae and background information are on file with author.

¹³⁸ 2005 S.D. Cir. LEXIS 1 (S.D. Cir. Ct. June 10, 2005).

¹³⁹ *Id.* at *7.

flower shop.¹⁴⁰ A Native American man entered the shop, identified himself as Sam Adams, and said that he was on his lunch break from Dakota Connection and wanted to buy some flowers for his girlfriend. As the customer started to write out an enclosure card at the counter, he asked Ms. Wanous to add something to his order.¹⁴¹ When Ms. Wanous went to a back workroom to get something, she was struck from behind on the head. When she awoke, she had no recollection of what had happened. After cleaning the blood from her hands and head, she immediately called the police upon noticing that her cash drawer was open and all the cash and checks had been taken.¹⁴² Shortly after Chief Flannery and Sergeant Fisher of the Sisseton Police Department arrived at the scene, Ms. Wanous was taken by ambulance to a hospital where she remained for five days after her injuries were discovered to include a four-inch cut to the back of her head, a fractured skull, and a concussion. Four days after being discharged from the hospital, Ms. Wanous provided the police with her initial statement regarding the incident, and several days later (two weeks after the incident), Ms. Wanous identified Adams from a photo lineup as the person who was in her store at the time she was attacked.¹⁴³ Defendant Sam Adams was subsequently prosecuted for aggravated assault, first-degree robbery, and first-degree burglary. A jury found Adams guilty of all three counts and he was sentenced to twenty-five years in prison.¹⁴⁴

The only physical evidence recovered from the scene that might link the defendant to the crime was the small enclosure card found on the counter that contained the hand printed phrase “To Karen From Sam.” No other physical evidence was recovered to connect Adams to the crime scene—no fingerprints, no blood, no DNA.¹⁴⁵

In granting Sam Adams’ petition for a writ of habeas corpus, the Circuit Court determined that the defense attorney’s laziness and complete incompetence undermined every aspect of Adams’ defense. The Court cited numerous instances of gross ineptitude on the part of the indigent defendant’s court-appointed attorney, with perhaps the most damaging one being his total lack of preparation concerning handwriting analysis.¹⁴⁶ With

¹⁴⁰ *Id.* at *1.

¹⁴¹ *Id.* at *2.

¹⁴² *Id.* at *3.

¹⁴³ *Id.* at *5–6.

¹⁴⁴ *Adams*, 2005 S.D. Cir. LEXIS 1 at *1–6, *30.

¹⁴⁵ *Id.*

¹⁴⁶ *Id.* at *44. Defense attorney Brankin failed to request, let alone consult with, an expert in forensic document examination, and stipulated to the State’s use of a handwriting expert two weeks before the start of the trial. “Brankin did not review Runyon’s report prior to trial, and had no knowledge of the technical standards used in handwriting analysis. . . . The skills and diligence Brankin exhibited in regards to this issue fell well outside the objective standard of reasonableness expected from competent trial counsel.” *Id.* at *44–45.

respect to the trial testimony of the State's handwriting expert, the Court noted:

The State's case against Adams was largely circumstantial in nature. The State did not produce any witnesses to the alleged crimes, other than the victim. The Sisseton Police Department did not collect any evidence that directly tied Adams to the scene of the crime. No fingerprints, weapon, or money was ever recovered. A major piece of physical evidence presented to the jury was the enclosure card left on Wanous' counter. As related earlier, it was inscribed with the words "To Karen, From Sam." However, the handwriting on the card was unnatural stick writing rather than normal printing. The State alleged this card was written by Adams. The defense maintained Adams did not write the card, and believed Sergeant Fisher forged it as evidence against Adams.

The State employed a forensic document examiner, Karen Runyon, as an expert to analyze the handwriting on the card. Brankin never attempted to employ his own handwriting expert, and never educated himself on the area of handwriting analysis. Approximately two weeks before the start of trial he stipulated to the use of Runyon as the handwriting expert without fully reviewing the contents of her report.

Prior to issuing her report, Runyon received writing exemplars of both Adams and Fisher that were analyzed against the card. Runyon stated that she received an inconclusive result when she analyzed Adams' handwriting, but that she could conclusively rule out Sergeant Fisher as the author. At trial, Runyon testified to her inconclusive finding in regards to Adams, but went on to detail similarities between Adams handwriting and the card. Brankin allowed her to give lengthy testimony on the issue without challenging her conclusions or prompting her to detail the similarities between the two.¹⁴⁷

Ms. Runyon testified that she had been employed as a FDE with the Minneapolis Police Department since 1978, and had also been accepting private sector civil casework assignments since 1988.¹⁴⁸ Her professional training included a four-year apprenticeship with the Questioned Document Unit of the Indiana State Police, FBI and US Secret Service training courses,

¹⁴⁷ *Id.* at *26–27.

¹⁴⁸ Transcript of Record at 381–86, *Adams v. Weber*, 2005 S.D. Cir. LEXIS 1 (S.D. Cir. Ct. 2001) (No. 03-107) (Karen Runyon's Testimony); Karen Runyon's Curriculum Vitae (on file with author).

and attendance at symposiums and workshops sponsored by professional membership organizations in the field.¹⁴⁹ She was a member of the Questioned Document Section of the AAFS and the Midwestern Association of Forensic Scientists (MAFS). Although not board certified, she had testified approximately 190 times in local, state, and federal courts.¹⁵⁰

At the habeas hearing, Adams presented evidence from three handwriting experts: Allan Keown, Vickie Willard, and Pat Girouard. All three experts, two of whom (Willard and Girouard) were Diplomates of the Board of Forensic Document Examiners (BFDE),¹⁵¹ echoed essentially the same concerns about the impropriety of comparing two sets of writings not suitable for comparison, and Ms. Runyon's bias in overstating an inconclusive opinion and providing disingenuous testimony.¹⁵² Each opined that Ms. Runyon was allowed to offer improper opinions that contravened the technical standards of handwriting analysis, and that Adams' defense attorney was not familiar with those standards and wholly unprepared to meaningfully challenge the admissibility of Runyon's opinions or to impeach and discredit her testimony. The two technical standards at issue were American Society for Testing and Materials (ASTM) standards, one establishing best practices for performing handwriting examinations and the other defining the standard terminology used by forensic document examiners in expressing conclusions. Willard pointed out that as an active member of ASTM Subcommittee E30.02 on Questioned Documents, Runyon had actually participated in writing and developing the two ASTM standards at issue.¹⁵³

The technical deficiencies of the handwriting opinions and biased nature of the trial testimony presented by Ms. Runyon (the State's expert) are specifically described below.

i. Incomparability of Writing Features: The Significance of
the Questioned Hand Printing Being Unnatural Stick
Printing and the Exemplars Being Natural Printing.

Ms. Runyon testified that the enclosure card found at the crime scene was "written in unnatural stick printing . . . as a means of disguise." Both Sam Adams' exemplars and Sergeant Fisher's exemplars were admittedly written in natural printing.¹⁵⁴

¹⁴⁹ *Id.*

¹⁵⁰ *Id.*

¹⁵¹ See *supra* note 126; Board of Forensic Document Examiners, <http://www.bfde.org/> (last visited Mar. 19, 2018).

¹⁵² Affidavit of Vickie L. Willard (Mar. 11, 2004) (on file with the author); Report of Allan Keown (Nov. 30, 2003) (on file with the author); Affidavit of Pat Girouard (Mar. 11, 2004) (on file with the author).

¹⁵³ Affidavit of Vickie L. Willard, *supra* note 152.

¹⁵⁴ Transcript of Record, *supra* note 148, at 416.

In comparing questioned writing consisting of unnatural stick printing with exemplars consisting only of natural printing, Ms. Runyon departed from the standard methodology and recognized best practices set forth in ASTM Standard E2290-03, Standard Guide for Examination of Handwritten Items (ASTM Standard E2290),¹⁵⁵ which provides:

§ 7.6.1: If [the questioned writing] is not natural writing, or . . . the available questioned writing is not suitable for comparison, discontinue these procedures and report accordingly.

§ 7.9.1: If [the known writing] is not natural writing, or . . . the available questioned writing is not suitable for comparison, discontinue these procedures and report accordingly.

§ 7.11.1: If the bodies of writing are not comparable, discontinue comparison and request comparable known writing, if appropriate.

ii. Distorting an Inconclusive Opinion in a Manner That
Favors a Particular Outcome

Ms. Runyon's conclusion as to whether Sam Adams wrote the unnatural stick printing on the enclosure card (marked at trial as "Exhibit 4") was stated to be "inconclusive", as indicated by the following excerpt from the official transcript of her direct testimony:

Q: When you did your comparison of Exhibit 4 to the items related to Sam Adams' handwriting, what was your conclusion from the comparison?

A: *My conclusion was that I was inconclusive.* There were both similarities to Mr. Adams' writing, as well as characteristics that were not found in the sample that I had of this writer. So based on the combination of what I had and what I did not have, I determined that with what was submitted that, actually, *a conclusion could not be rendered in one direction or another.*¹⁵⁶

¹⁵⁵ ASTM INTERNATIONAL, ASTM STANDARD E2290-03: STANDARD GUIDE FOR EXAMINATION OF HANDWRITTEN ITEMS 3 (2003).

¹⁵⁶ Transcript of Record, *supra* note 148, at 392 (emphasis added).

ASTM Standard E1658-96, Standard Terminology for Expressing Conclusions of Forensic Document Examiners (ASTM Standard E1658),¹⁵⁷ recommends and defines several terms that FDEs should use to express the level of confidence associated with their opinion(s); it provides a standardized framework for understanding the true meaning of the level of confidence associated with an opinion or conclusion expressed by a FDE. As defined in ASTM Standard E1658, the terms “inconclusive” and “indeterminable” are synonymous and represent “the zero point of the confidence scale”; these terms are “used when there are significant limiting factors, such as disguise in the questioned and/or known writing or a lack of comparable writing, and the examiner does not have a leaning one way or the other.”¹⁵⁸

Once Ms. Runyon expressed an inconclusive opinion as to whether Sam Adams wrote the card and testified that no conclusion could be reached one way or the other, the only proper and accurate statement that she could make was that Sam Adams “cannot be eliminated or identified as the writer.” However, Ms. Runyon chose to embellish her testimony with an inaccurate and misleading statement designed to support the prosecution-favored outcome, as reflected in the following exchange during her direct testimony:

Q: Ms. Runyon, you’re not saying that Sam did not write this card?

A: No. *You cannot eliminate him as a writer*, no.

Q: But you’re not saying that he did?

A: *Neither can you identify him positively as the writer*, no.¹⁵⁹
(Emphasis added)

Ms. Runyon’s gratuitous inclusion of the word “positively” in testifying that Sam Adams *cannot be eliminated or positively identified* as the writer clearly manifested bias in favor of the prosecution. This overstatement was highly prejudicial to the defendant because it wrongfully implied a “near match,” i.e., that the defendant *can be identified as the writer, but just not positively*. This form of disingenuous testimony is not uncommon in criminal cases involving prosecution handwriting experts who appear to be unduly influenced or motivated to testify in a manner that

¹⁵⁷ ASTM INTERNATIONAL, ASTM STANDARD E1658-96: STANDARD TERMINOLOGY FOR EXPRESSING CONCLUSIONS OF FORENSIC DOCUMENT EXAMINERS (1996).

¹⁵⁸ *Id.* at 2.

¹⁵⁹ ASTM STANDARD E1658-96, *supra* note 157, at 1.

suggests support for the inculpatory hypothesis even when the evidence itself favors neither the inculpatory nor exculpatory hypothesis.

iii. Misinterpreting Evidence or Providing Exaggerated
Testimony to Support a Favored Outcome

The defense maintained that Adams did not write the card, and suggested that Officer Fisher fabricated it as evidence against Adams.¹⁶⁰ The following excerpt of Runyon's trial testimony concerns the results of her examination and comparison of the unnatural stick printing on the enclosure card with the naturally written hand printing exemplars of Officer Fisher:

Q: What was your conclusion *from the comparison of Officer Fisher's known handwriting* to the questioned document?

A: My conclusion was that it was *highly probable that he was not the writer of the questioned material*.¹⁶¹ (Emphasis added)

ASTM Standard E1658, *supra*, defines "highly probable" as meaning that "the evidence is very persuasive" and "the examiner is virtually certain" of the conclusion (opinion) expressed.¹⁶² Hence, Ms. Runyon concluded that Officer Fisher could be eliminated with virtual certainty as the writer of the unnatural stick printing appearing on the enclosure card.

As noted earlier, no conclusion could be rendered in one direction or another regarding authorship of the questioned writing because the unnatural stick printing appearing on the card was not suitable for comparison with the naturally written exemplars available for both Adams and Officer Fisher. Accordingly, an inconclusive opinion was warranted with respect to whether Officer Fisher wrote the enclosure card for the same reason Ms. Runyon reached an inconclusive opinion with respect to whether defendant Adams wrote the card – unnatural stick printing cannot be compared to natural printing.

The only way Officer Fisher could have properly been eliminated as the writer of the enclosure card was by evidence showing that his writing skills were so impaired as to have made it impossible for him to produce the

¹⁶⁰ The Circuit Court pointed out inconsistencies in the trial testimony of the two police officers that responded to the crime scene. "Chief Flannery and the victim both testified the victim did not give the officers a physical description of the perpetrator. Sergeant Fisher stated the victim did provide a description. The two officers testified inconsistently as to who arrived before whom at the Treasure Chest. They both testified that the other was the officer to actually pick up the enclosure card from Wanous' counter. They both stated that they considered the other officer to be in charge of the investigation." *Adams v. Weber*, No. 03-107, 2005 S.D. Cir. LEXIS 1, at *34-35 (S.D. Cir. Ct. June 10, 2005).

¹⁶¹ See *supra* note 148, at 415.

¹⁶² See *supra* note 157, at 1, § 4.1.

unnatural stick printing at issue. Such evidence being absent, the inclination on the part of the prosecution's handwriting expert to disassociate the disguised hand printing on the enclosure card from the natural hand printing of Officer Fisher clearly reflects a biased conclusion derived from an illogical interpretation of evidence, presumably resulting from Ms. Runyon's adoption of an adversarial role in which the outcome-oriented goal trumped the truth-seeking goal.

3. Improper Methodology—Relying Solely Upon *Post Litem Motem* Signatures Supplied by A Disclaiming Signatory to Support an Opinion of Forgery (the *Almeciga* case)

Returning to *Almeciga v. Center for Investigative Reporting*, discussed in regard to the role of verification, *supra*, the plaintiff in the case denied having signed or even seen the key document, a Release.¹⁶³ As previously noted, the putative expert called by the defense was Wendy Carlson. Ms. Carlson had compared the contested signature to several "known" examples provided by the plaintiff and concluded that the contested signature was a forgery. All the known samples Carlson used were signed well after the litigation commenced (*post litem motem* exemplars). Following a *Daubert* hearing, U.S. District Court Judge Jed. S. Rakoff of the Southern District of New York excluded the proffered testimony of plaintiff's handwriting expert, Wendy Carlson, finding that her testimony was far too problematic to be admissible under Rule 702 as technical or otherwise 'specialized' expert testimony, even on a *Kumho Tire* approach,"¹⁶⁴ and that her "methodology [was] fundamentally unreliable and critically flawed in so many respects [that her] testimony would be more likely to obfuscate the issues in the case than to 'help the trier of fact to understand the evidence or to determine an issue.'"¹⁶⁵

Judge Rakoff cited the following four factors as contributing to his conclusion that Carlson's handwriting opinion was unreliable and inadmissible:

- 1) Bias (Carlson knew the favored outcome from the start)¹⁶⁶
- 2) "The subjectivity and vagueness of Carlson's analysis severely diminishes the reliability of Carlson's

¹⁶³ *Almeciga v. Ctr. for Investigative Reporting, Inc.*, 185 F. Supp. 3d 401, 409 (S.D.N.Y. 2016).

¹⁶⁴ *Id.* at 424.

¹⁶⁵ *Id.* at 426.

¹⁶⁶ Judge Rakoff pointed out how the attorney who retained Ms. Carlson had engaged in "blatant biasing tactics [that] compromised Ms. Carlson's ability to provide a neutral examination, a danger made even greater by the highly subjective nature of Carlson's methodology." *Id.* at 425.

methodology.”¹⁶⁷

(3) Carlson’s analysis is “critically flawed” because she failed to consider the possibility of disguise—that the disclaiming signatory provided disguised signature exemplars when the evidence clearly suggested that to be the case;¹⁶⁸ and

(4) Striking contradictions between Carlson’s reports and in-court testimony “diminished Carlson’s credibility”¹⁶⁹ such as:

(a) Carlson claimed to use ACE-V, but due to time constraints, omitted the V (Verification),¹⁷⁰ discussed in Part V.B, *supra*.

(b) Carlson’s initial report stated that “the signature on the Release was ‘made to resemble’ Plaintiffs,” but Carlson’s in-court testimony acknowledged that the Release signature and the known signatures “weren’t even close” and that the Release signature “was not like an attempted forgery.”

Of the above four factors, only Ms. Carlson’s use of the self-serving *post litem motem* exemplars supplied by the disclaiming signatory was expressly stated by Judge Rakoff to have rendered Ms. Carlson’s analysis “critically flawed.” As indicated *supra*, this departure from standard methodology is acknowledged to be one of the principle sources of error in opinions declaring a signature to be forgery. Hence, this factor alone sufficed to render Ms. Carlson’s opinion unreliable and inadmissible under Rule 702, especially given the fact that Judge Rakoff’s own inquiry led him to view documents filed in other state court proceedings that contained *ante litem motem* exemplar signatures of the plaintiff which he observed to be remarkably similar to the contested signature she denied having signed or even seen.¹⁷¹ Moreover, Judge Rakoff properly inferred that Ms. Carlson

¹⁶⁷ *Almeciga*, 185 F. Supp. 3d at 426.

¹⁶⁸ *Id.*

¹⁶⁹ *Id.*

¹⁷⁰ *Id.*

¹⁷¹ “[W]hile testimony that accounted for the possibility of disguise and addressed why the ‘known’ signatures were not the product of intentional disguise could at least have potentially assisted the trier of fact, Carlson did not offer such testimony. To the contrary, Carlson confirmed at her deposition that she was ‘relying on the plaintiff’s representations that [the known signatures] are *accurate* representations of her signature.’ This is a critical flaw in Carlson’s methodology because it assumes away a key issue: whether *Almeciga* intentionally disguised her handwriting in producing the known samples after this dispute was initiated or whether the known samples accurately represent her actual handwriting. By relying on plaintiff’s counsel’s representation that the ‘known’ signatures were accurate representations of plaintiff’s signature, the result of Carlson’s analysis was effectively pre-ordained and her testimony cannot be considered the ‘product of reliable principles and methods.’ In fact, Carlson’s testimony has been excluded by at least one other court in part on such a basis. See *United States v. LeBeau*, 2015 WL 4068158, at *8 (D.S.D. June 10, 2015) (‘[Carlson’s] analysis and opinions entirely hinge on whether she received an accurate ‘known’ signature from [the defendant].’). The tainting effect of Carlson’s assumption in this

had to have been aware of the distinct possibility that Ms. Almeciga supplied her with disguised exemplar signatures because Ms. Carlson's in-court testimony acknowledged that the contested signature and the exemplar (known) signatures "weren't even close" and that the contested Release signature "was not like an attempted forgery."¹⁷²

In finding that "handwriting analysis is *unlikely* to meet the admissibility requirements of Federal Rule of Evidence 702, and that, *in any event*, Ms. Carlson's testimony does not meet these standards," Judge Rakoff's use of the term "unlikely" implies something other than a definitive finding, whereas the use of the phrase "in any event" indicates that what follows constitutes the basis for his decision to exclude Ms. Carlson's testimony. This is consistent with the opinion of U.S. District Court Judge Martinez in *Crew Tile Distribution*, which "[read] the exclusion of Carlson's testimony in *Almeciga* as resting on a case-specific analysis under *Kumho Tire*," and not on a general disqualification of forensic handwriting analysis under *Daubert*.¹⁷³

VII. CONSIDERATIONS AND SUGGESTIONS FOR PROCEEDING FURTHER ALONG "THE PATH FORWARD"

A. *What the Community of FDE Practitioners Needs to Do*

1. Restate the Foundational Principles Supporting the Discipline of Forensic Handwriting Analysis as a Technical Skill

Unlike the Latent Print Examination community, the FDE community has been slow to abandon the claim of uniqueness as a foundational basis for forensic handwriting analysis. Despite recent court decisions pointing out that uniqueness is something that has not been and probably never can be validated by empirical proof,¹⁷⁴ and that it is not a prerequisite for the admissibility of handwriting identification opinion evidence as a technical skill under Rule 702, the FDE community has yet to come to terms with the

regard may be gleaned from what she infers on the basis of her observation that the 'signature on the questioned document is written with great fluidity and a faster speed, unlike the known signatures that display a slower, more methodical and unrefined style of writing.' To Carlson, who took on faith that the 'known' signatures were accurate representations of plaintiff's handwriting, this discrepancy is evidence that the Release was forged. Yet, at the Daubert hearing, Carlson confirmed that slower, methodical handwriting was equally consistent . . . or maybe even more consistent with someone trying to fake the known signatures[.]'" *Almeciga*, 185 F. Supp. 3d at 425–26 (citations omitted).

¹⁷² *Almeciga*, 185 F. Supp. 3d at 426.

¹⁷³ *Crew Tile Distribution, Inc. v. Porcelanosa L.A., Inc.*, No. 13-cv-3206, 2016 WL 8608447, at *9 (D. Colo. Sept. 12, 2016).

¹⁷⁴ See discussion of *U.S. v. Kelly* and *U.S. v. Johnsted* in *supra* Part II.

need to restate the foundational principles supporting the discipline of forensic handwriting analysis.

The theory of uniqueness needs to be replaced with a neuroscience model based upon motor control behavior and complexity theory, along the lines discussed in Part III of this Article. Practitioners must abandon the practice of characterizing the discipline of forensic handwriting analysis as a “science,” or describing handwriting examinations as “scientific.” FDEs need to acknowledge the subjectivity of forensic handwriting analysis and understand that the subjectivity of a claimed expertise does not preclude a court from finding that such expertise is sufficiently reliable as a technical skill to be admitted as opinion evidence under Rule 702, as under *Daubert* and *Kumho Tire*.

2. Develop More Robust and Transparent Best Practice
Standards for Examining Handwritten Items and Expressing
Opinions in Reports and Testimony

The FDE community has been more active than any other forensic discipline in producing professional standards, having published an array of twenty-one standards through ASTM International, a private consensus standards development organization (SDO).¹⁷⁵ Most of those standards were initially drafted by the FBI-funded Scientific Working Group for Forensic Document Examination (SWGDOC) and thereafter were revised and voted upon by members of ASTM’s Questioned Documents Subcommittee E30.02, one of several discipline-specific subcommittees operating under the umbrella of ASTM Main Committee E-30 on Forensic Science.¹⁷⁶ Unfortunately, the content of some of those standards was not without its critics, and the criteria used by E30.02 officers in classifying voting interests resulted in the filing of formal complaints and appeals alleging that the E30.02 Questioned Document subcommittee’s method of classifying voting interests produced an imbalance in stakeholder interests that contravened the SDO’s essential due process requirements.¹⁷⁷ In April 2012, confronted with allegations of due process violations and of improperly using the ASTM standards development process to obtain an unfair economic advantage in the marketplace of forensic document examination services, the coalition of public sector subcommittee members opted to use their supermajority control of voting to shut down Subcommittee E30.02 and return to using the

¹⁷⁵ See ASTM Int’l, *Subcommittee E30.90 on Executive*, <https://www.astm.org/COMMIT/SUBCOMMIT/E3090.htm> (last visited Mar. 19, 2018).

¹⁷⁶ See ASTM Int’l, *Standards Formerly Under the Jurisdiction of E30.02 on Questioned Documents*, <https://www.astm.org/COMMIT/forensic-docs/index.html> (last visited Mar. 19, 2018).

¹⁷⁷ Letter from Andrew Sulner to Steve Orthey, Sec’y, ASTM Committee of Technical Committee Operations (COTCO) (July 14, 2009) (on file with author).

SWGDOC group to generate best practice standards for the FDE profession.¹⁷⁸

Since best practice standards and guidelines produced by professional organizations or Scientific Working Groups (SWGs) do not meet the requirements of the National Technology Transfer and Advancement Act (NTTAA)¹⁷⁹ for being voluntary consensus standards, the American Academy of Forensic Sciences (AAFS) established in 2016 an SDO called the AAFS Academy Standards Board (ASB) to develop voluntary consensus standards, technical reports and best practice recommendations for forensic science disciplines. The ASB has been accredited by the American National Standards Institute (ANSI), which requires that all SDOs adhere to certain essential requirements to ensure that its procedures for developing voluntary consensus standards adhere to principles of fundamental fairness. These requirements include openness (transparency), balance of stakeholder interests, due process, and a meaningful appeals process.¹⁸⁰ Perhaps the biggest challenge facing the ASB will be to ensure that its forensic standards development process is not dominated by a single stakeholder interest and that the standards it promulgates do not conflict with standards proposed by NIST, which formed the Organization of Scientific Area Committees (OSAC) for Forensic Science “to strengthen the nation’s use of forensic science by facilitating the development of technically sound forensic science

¹⁷⁸ Group-serving bias, characterized by coordinated action towards a common goal by members of a specific group or coalition, is yet another source of bias that impacts all sorts of decision-making within the forensic science community. It is one of the dangers and risks of having the process for developing forensic standards controlled by members of the law enforcement community or any other coalitional alliance, especially when it comes to the development of standards recommending minimum training requirements for practitioners such as forensic document examiners. This became apparent with the closure of ASTM Subcommittee E30.02 on Questioned Documents, which ensued after private sector voting members of E30.02 and other interested stakeholders filed appeals and registered complaints alleging that a coalition of E30.02 subcommittee members affiliated with government agencies and government-sponsored membership organizations were violating ASTM due process requirements and federal antitrust laws prohibiting unfair trade practices. See Andrew Sulner, *A Critical Look at Some Needed Reforms in the Landscape of Forensic Science Education and Mentorship Training Standards*, 24 J. FORENSIC DOCUMENT EXAMINATION 73 (2014); Vickie Willard, *ASTM Committee E30 on Forensic Sciences: The History of Subcommittee E30.02 on Questioned Documents*, ASS’N FORENSIC DOCUMENT EXAMINERS (2013), <http://afde.org/resources/The-History-of-E30-02.pdf>.

¹⁷⁹ National Technology Transfer and Advancement Act of 1995 (NTTAA), Pub. L. No. 104-113, 110 Stat. 775 (1996), directs that except where inconsistent with applicable law or otherwise impractical, all federal agencies and departments must use technical standards that are developed or adopted by voluntary consensus standards bodies, in lieu of creating proprietary, non-consensus “in-house” standards.

¹⁸⁰ See Am. Acad. of Forensic Sci., *Why Do We Need an SDO If We Have the OSAC?*, ACAD. NEWS (Aug. 4, 2016), <https://news.aafs.org/asb-news/why-do-we-need-an-sdo-if-we-have-the-osac/>; *AAFS Standards Board*, <https://asb.aafs.org> (last visited Mar. 19, 2018).

standards and by promoting the adoption of those standards by the forensic science community.”¹⁸¹

i. Handwriting Examinations: The Need for Standards to
Incorporate Practical Measures Designed to Reduce the
Potential for Cognitive Bias

Reducing the Likelihood that Cognitive Bias Will Taint the Examiner

ASTM E2290, *Standard Guide for the Examination of Handwritten Items*, was the only consensus standard adopted by the FDE profession that set forth best practice guidelines for performing forensic handwriting examinations.¹⁸² First published in 2003, it was reviewed once thereafter in 2007, and republished by ASTM without any substantive change from the 2003 version. E2290 expired without renewal upon the closure of subcommittee E30.02 in 2012, but it has since been republished as *SWGDOC Standard for Examination of Handwritten Items*, which contains the same language as the E2290 standard, apparently with the permission of ASTM International. Each of these published versions of a *Standard Guide for the Examination of Handwritten Items* fails to mention, let alone suggest, the implementation of protocols designed to reduce potentially biasing influences, especially contextual bias.¹⁸³

The FDE community’s resistance to implementing bias control measures as standard protocol may be attributable to the fact that there are still far too many FDEs who remain adamant in their belief that everything is relevant to the tasks they perform, or that proper training and experience, or board certification, renders them immune from being influenced by task-irrelevant case information. Or, it may simply result from the natural tendency of human beings to take comfort in knowing that their subjective opinions (“expert” or not) are confirmed by independent facts they believe to be true at the time, even though those facts may subsequently prove to be false, such as an eyewitness account or a defendant’s confession.

At the very least, I believe the concept of sequential unmasking needs to be accepted and implemented as a standard protocol for forensic handwriting investigations, with guidelines established for identifying what types of

¹⁸¹ *The Organization of Scientific Area Committees for Forensic Science*, NAT’L INST. OF STANDARDS & TECH. (Apr. 11, 2016), <https://www.nist.gov/topics/forensic-science/organization-scientific-area-committees-osac> (last updated Mar. 13, 2018).

¹⁸² See ASTM Int’l, *ASTM E2290-07a: Standard Guide for Examination of Handwritten Items* (2007), <https://www.astm.org/Standards/E2290.htm>; Sci. Working Grp. for Forensic Document Examination, *SWGDOC Standard for Examination of Handwritten Items: Version 2013-1* (Jan. 14, 2015), https://www.nist.gov/sites/default/files/documents/2016/10/26/swgdoc_standard_for_examination_of_handwritten_items.pdf.

¹⁸³ *Id.*

information might be deemed to be domain-relevant, and at what stage of the evidence evaluation process such information can be properly disclosed to the analyst.¹⁸⁴ Additionally, FDEs working in the private sector must be encouraged to advise and forewarn attorneys and clients at the outset of the initial communication discussing their possible retention as an expert, not to disclose or imply what the favored outcome is, explaining why such “blinding” would be preferable and ultimately in the client’s best interest. Examiners working in both the public and private sectors should also make every effort to use “blind testing” procedures to further facilitate an objective analysis of the evidence and thereby effectively foreclose an opposing attorney’s ability to use contextual bias as a means of impeaching otherwise credible results and opinions.

“Blind Testing” Protocol Used to Examine Hand Printing in United States v. Matusiewicz

Most recently, the benefits of using blind testing and sequential unmasking procedures in performing forensic handwriting examinations were expressly acknowledged by both the defense attorneys and the prosecutors involved in *United States v. Matusiewicz*, the nation’s first federal prosecution of defendants charged with “cyberstalking resulting in death.”¹⁸⁵ The defendants were a mother (Lenore), son (David), and daughter (Christine), who were indicted under the new federal anti-cyberstalking statute after the father (Tom) gunned down the son’s ex-wife and her friend in a Wilmington, Delaware courthouse lobby in 2013 while walking into a child support hearing, before taking his own life. I was retained as a handwriting expert in this case by Edson A. Bostic, Chief Federal Public Defender for the District of Delaware, who was representing the son, David Matusiewicz. At the very outset, I advised Mr. Bostic of the need to avoid letting me know what his favored outcome is and that, to the extent feasible, I be supplied with anonymous hand printing exemplars of several individuals for purposes of comparing same to the questioned hand printing. Accordingly, I was provided with digital photographs of four documents bearing questioned hand printing (items I designated as Q-1 through Q-4) and scanned images of documents submitted as bearing hand printing exemplars of four writers identified simply as Writers A, B, C, and D.¹⁸⁶ Based upon my examination of these items, I issued a report dated

¹⁸⁴ See Bryan Found & John Ganas, *The Management of Domain Irrelevant Context Information in Forensic Handwriting Examination Casework*, 53 SCI. & JUST. 154–58 (2013), and authorities cited in note 1, *supra*.

¹⁸⁵ See Lauren Walker, *Family Receives Life in Prison for First-Ever Cyberstalking Conviction*, NEWSWEEK (Feb. 26, 2016), <http://www.newsweek.com/family-receives-life-prison-first-ever-cyberstalking-conviction-430833>.

¹⁸⁶ To preserve anonymity of the exemplars, each document bearing known hand printing

2018]

A DOCUMENT EXAMINER TELLS ALL

705

March 16, 2015 which set forth the basis for my opinions that the hand printed entries appearing on item Q-1 were written by Writer C and the hand printed entries appearing on items Q-2 through Q-4 were written by Writer B.

Thereafter, I received a copy of an FBI Laboratory Report of Examination dated January 27, 2015, prepared and issued by Daniel P. Anderson of the Questioned Documents Unit of the FBI Laboratory in Quantico, Virginia. Upon reading this report, I was shocked to discover that Mr. Anderson had reached an inconclusive opinion with respect to who authored items Q-2 through Q-4, despite having examined the same documents that I had examined, albeit without using a “blind testing” protocol.¹⁸⁷ This was also when I first learned the identities of Writers A, B, C and D, because FDE Anderson was provided with additional known writing consisting of compulsory (request) exemplars of hand printing specifically identified as having been obtained from David Matusiewicz (the son), Lenore Matusiewicz (the mother), and Amy Gonzalez (the daughter), as well as additional documents bearing known hand printing exemplars attributable to Tom Matusiewicz (the deceased father and “shooter”). I was also curious why six other items submitted to Mr. Anderson as bearing questioned hand printing were never examined by him, and why the document identified as Q-1 in my March 16, 2015 report was not submitted to Anderson (and hence, not referenced in his report). It was then that I learned that the Government was alleging that the “HL” hand printed at the top of Q-1 stands for “Hit List,” and the hand printed names appearing underneath the “HL” included the name of the first federal judge assigned to the case, who recused himself on that account.

I subsequently examined all documents submitted to FBI examiner Anderson that I had not seen before, as well as higher resolution color flatbed-scanned images of documents that I had previously received in the form of photos taken with a digital camera. Based upon my examination of these additional items, I prepared a second report dated June 16, 2015, wherein I confirmed my conclusion that Writer B, now identified as Lenore Matusiewicz, authored items Q-2 through Q-4, and that Writer C, now identified as Tom Matusiewicz, authored item Q-1 as well as the six hand printed items (designated by me as Q-5 through Q-10) that were submitted to Anderson but never examined by him.

During my testimony, I used numerous illustrative charts to aid the jury in observing the hand printing features that supported my opinions, and stated, “quite frankly . . . any expert with basic elementary training would

was redacted to remove any information that could identify the writer.

¹⁸⁷ The FBI Lab Report failed to provide any explanation or reason(s) why the handwriting analyst (Anderson) concluded that the hand printing evidence submitted for analysis was insufficient to render an opinion.

have reached the same conclusion I did.”¹⁸⁸ My expertise and opinions were never challenged by the Government. But even more remarkable was the following exchange that took place in the presence of the jury at the end of my very brief cross-examination by Assistant United States Attorney Edward J. McAndrew, a federal prosecutor with extensive trial experience in high profile cases who, at the time, was the Justice Department’s Cybercrime Coordinator and national security cybercrime specialist:

Q. All right. And my last question for you, Mr. Sulner, is should we need to hire you in the future, can we reach you at this address?

A. Yes, sir.

Mr. McAndrew: Thank you. Thank you, Your Honor.¹⁸⁹

After the case was concluded, Mr. Bostic wrote me a letter in which he indicated his unequivocal support for the blind testing/sequential unmasking protocol I utilized:

I thank you for the excellent work you did on my client’s behalf in the above captioned matter. Over my 30 years of litigation, I have worked with numerous experts, including other forensic document examiners, architects and engineers, medical doctors, psychologists, accident reconstruction experts, arson investigators, construction management consultants, and securities and other fraud experts, to name a few. Your commitment, professionalism and thoroughness were at the highest level of any expert with whom I have ever worked. The most significant and impressive contribution you made to our case by far, however, was your introduction of blind testing. This was my first exposure to blind testing. Your insistence upon utilizing this methodology clearly carried the day with respect to the reliability of the findings and conclusions you reached and testified to during trial. As you may recall, after presenting your direct testimony, the government stood only to concede and acknowledge the accuracy and significance of your findings, while noting that they may want to use your services and/or your

¹⁸⁸ Transcript of Trial Testimony at 5195, *United States v. Matusiewicz*, 165 F. Supp. 3d 166 (D. Del. 2015) (on file with author).

¹⁸⁹ *Id.* at 5200. During the court recess that immediately followed my testimony, Mr. McAndrew informed me that he asked me the last question because he wanted the record to reflect that he and his fellow prosecutors disagreed with the inconclusive opinion(s) provided by the FBI’s document examiner and considered my identification of the two writers to be reliable and accurate.

2018]

A DOCUMENT EXAMINER TELLS ALL

707

blind testing procedures in other similar matters going forward. Of all the experts that we used or sought to admit at trial, your report and testimony was the only one which the Government did not challenge. Such concession, on critical documents in evidence in this case, was essential to our defense theory and particularly noteworthy given the FBI's inconclusive findings based upon their expert's examination of the very same evidence you examined and about which you testified. Therefore, I thank you again for your hard work and dedication to my client's case. . . .¹⁹⁰

ii. Handwriting Examinations: The Need to Delineate the Decision-Making Stages Involved in the Examination and Evaluation Process

Specifying ("Mapping Out") the Steps Involved in the Examination and Evaluation Process

Both ASTM E2290 and its SWGDOC counterpart were too vague in describing the procedures involved in the forensic examination of handwritten items; they failed to address the various critical decisions that are commonly made by examiners in the examination and interpretive phases of performing forensic handwriting examinations. In 1999, the Journal of Forensic Document Examination published a Modular Method for Examining Handwriting (The Modular Method) that formed the framework of the forensic handwriting methodology used by government laboratories in Australia and New Zealand, as recommended by the Document Examination Specialist Advisory Group (DocSAG) of Australia and New Zealand.¹⁹¹ The Modular Method provided a series of modules describing the nature of each decision to be made by an examiner at each stage of the examination process, with a flow chart summarizing the decision-making stages within each module. SWGDOC considered adopting this type of modular approach in 2000 when it submitted its initial draft of a proposed standard for examining handwritten items. However, that idea was abandoned and has never been seriously reconsidered until recently, undoubtedly precipitated by the invaluable input provided by the late Dr. Bryan Found while a member of the NIST-OSAC Expert Working Group on Human Factors in Handwriting Examination, and the publication in the

¹⁹⁰ Letter from Edson A. Bostic, Del. Fed. Pub. Defender, to author (Sept. 11, 2015) (on file with author).

¹⁹¹ Bryan Found & Doug Rogers, *Documentation of Forensic Handwriting Comparison and Identification Method: A Modular Approach*, 12 J. FORENSIC DOCUMENT EXAMINATION 1 (1999).

Journal of Forensic Document Examination of *The Modular Forensic Handwriting Method (2016 Version)*, an extensive update of the prior 1999 version.¹⁹²

The Modular Forensic Handwriting Method (2016 Version), developed by Found, Carolyne Bird, and their colleagues, consists of a series of ten modules, the first of which is “The Method Flow Diagram,” a detailed process map which summarizes the essential decision-making points and options within the handwriting examination/evaluation process.¹⁹³ Modules 2 through 10 provide detailed descriptions and graphic illustrations for each of the evaluative and interpretive steps that a forensic handwriting analyst needs to perform in reaching an opinion or conclusion, including an opinion that the handwriting evidence is insufficient to warrant any conclusion. These nine modules cover the following topic areas:

- Module 2: Determination of the suitability of questioned and comparison samples: issues of comparability and contamination
- Module 3: Comparison of handwriting samples
- Module 4: Non-original handwriting
- Module 5: The assessment of handwriting complexity: addressing propositions regarding authorship
- Module 6: Structural and line quality dissimilarities
- Module 7: Traced writings
- Module 8: Unnatural handwriting behaviors
- Module 9: Line quality and skill
- Module 10: Evidence evaluation and reporting procedures

The Modular Forensic Handwriting Method (2016 Version) should serve as a model for developing a handwriting examination standard that is more rigorous and transparent than what the American community of FDEs has developed to date, especially since Found and Bird reported positive findings from a multi-year study of over 27,000 opinions expressed by 28 document examiners designed to test the validity of their modular approach.¹⁹⁴ By requiring examiners to document their decisions for each decision point along the continuum of tasks involved in the examination and interpretive stages of the process, the modular approach affords a more effective and transparent means of evaluating the reliability of the key decisions that purportedly support an examiner’s conclusion. For example, one decision point of concern to the FDE community as well as the courts

¹⁹² Bryan Found & Carolyne Bird, *The Modular Forensic Handwriting Method 2016 Version*, 26 J. FORENSIC DOCUMENT EXAMINATION 7 (2016).

¹⁹³ *Id.*

¹⁹⁴ *Id.* at 63–70.

concerns an examiner's decision as to whether perceptible differences between two sets of writings should be attributable to natural variation (intra-writer variability), different writers (inter-writer variability), or disguise (intentional distortion), especially in view of research indicating that presumptively well-trained FDEs have a high error rate when performing tasks of this nature.¹⁹⁵

Developing a robust and transparent American national handwriting examination standard along the lines of *The Modular Forensic Handwriting Method (2016 Version)* would serve to alleviate some of the criticisms directed at the subjectivity of forensic handwriting examination as well as improve the accuracy and strengthen the probative value of handwriting identification opinions expressed by trained examiners.

iii. Handwriting Identification Opinions: The Need to Establish
More Meaningful and Appropriate Language for
Expressing Qualitative Levels of Confidence (Certitude)

The problems associated with document examiners using different terminology in reporting the results of handwriting comparisons was brought to the forefront in a series of articles written by McAlexander, Ellen, and Purdy in the late 1970's and early 1980's.¹⁹⁶ Their suggested language for use in expressing handwriting identification opinions was later presented in a 1991 article published as a Letter to the Editor in the *Journal of Forensic Sciences* that proposed the use of a nine-level scale comprising terms for expressing subjective (non-quantified) estimates of probability with respect to handwriting identification opinions.¹⁹⁷ The 1991 article was adopted as recommended guidelines in reports and testimony by the Questioned Document Section of the AAFS and the ABFDE.¹⁹⁸ The nine-level scale consisted of a positive category of four rank-ordered expressions of confidence level—identification (definitely did write), strong probability did write, probably did write, and indications (evidence to suggest) may have written; a neutral category—inconclusive (no conclusion); and a negative category of the corresponding four rank-ordered levels of confidence—

¹⁹⁵ See D. Michael Risinger, *Cases Involving the Reliability of Handwriting Identification Expertise Since the Decision in Daubert*, 43 TULSA L. REV. 477, 549 (2008).

¹⁹⁶ Thomas V. McAlexander, *The Meaning of Handwriting Opinions*, 5 J. POLICE SCI. & ADMIN. 43 (1977); D.M. Ellen, *The Expression of Conclusions in Handwriting Examination*, 12 CANADIAN SOC'Y FORENSIC SCI. J. 117 (1979); Dan C. Purdy, *The Requirements of Effective Report Writing for Document Examiners*, 15 CANADIAN SOC'Y FORENSIC SCI. J. 146 (1982).

¹⁹⁷ Thomas V. McAlexander, Jan Beck & Ronald M. Dick, *The Standardization of Handwriting Opinion Terminology*, 36 J. FORENSIC SCI. 311, 314–16 (1991).

¹⁹⁸ *Id.* at 311. These guidelines were later incorporated in a consensus standard first published in 1996 as ASTM Standard E-1658, *Standard Terminology for Expressing Conclusions of Forensic Document Examiners*.

elimination (definitely did not write), strong probability did not write, probably did not write, and indications (evidence to suggest) may not have written.

On February 28, 1995, ABFDE Diplomate Mary Wenderoth Kelly testified at a *Daubert* hearing in *United States v. Starzecpyzel*¹⁹⁹ about the FDE profession's adoption of the nine-level scale as the standard terminology to be used by FDEs in reporting the results of handwriting examinations.²⁰⁰ United States District Court Judge McKenna of the Southern District of New York was unimpressed; he rejected the nine-level scale, finding that it is highly subjective and imprecise and can easily lead to misleading and improper testimony.²⁰¹

Undaunted by Judge McKenna's April 3, 1995 decision in *Starzecpyzel*, on September 10, 1995, ASTM Subcommittee E30.02 on Questioned Documents voted to approve the nine-level scale as a consensus standard for expressing handwriting identification opinions, and it was officially published in April 1996 as ASTM Standard E1658-05, *Standard Terminology for Expressing Conclusions of Forensic Document Examiners*. Subsequently, the original 1996 version of E1658 was reviewed and republished without any substantive changes in 1997, 2004 and 2008. From 1996 until now, ASTM E1658 has been routinely cited by FDEs as the profession's consensus standard for expressing handwriting identification opinions in reports and testimony.

When I first read Judge McKenna's decision in *Starzecpyzel*, I was relieved to see that a prominent jurist shared the same concerns about vagueness and unreliability that I had expressed in negative votes and comments whenever ASTM E1658 came up for discussion and review as an ASTM standard. What was most troubling to me as a lawyer and a forensic document examiner was that ASTM E1658 sanctioned testimony "that there is evidence which *indicates* (or *suggests*) that the John Doe of the known material may have written the questioned materials but the evidence falls far short of that necessary to support a definite conclusion"²⁰² The "Discussion" paragraph inserted as a note immediately following the definition of "indications" in each and every published version of the ASTM E1658 standard contained the following caveat:

This is a very weak opinion, and a report may be misinterpreted to be an identification by some readers if the report simply states, "The evidence *indicates* that the John Doe of the known material wrote the questioned material." There should always be additional

¹⁹⁹ *United States v. Starzecpyzel*, 880 F. Supp. 1027 (S.D.N.Y. 1995).

²⁰⁰ *Id.* at 1028.

²⁰¹ *Id.* at 1048.

²⁰² ASTM INT'L, ASTM E1658-08 2 § 4.1(2008).

limiting words or phrases (such as “may have” or “but the evidence is far from conclusive”) when this opinion is reported, to ensure that the reader understands that the opinion is weak. Some examiners doubt the desirability of reporting an opinion this vague, and certainly they cannot be criticized if they eliminate this terminology. But those examiners who are trying to encompass the entire “gray scale” of degrees of confidence may wish to use this or a similar term.²⁰³

Unfortunately, my FDE colleagues in ASTM failed to appreciate my objection that this type of handwriting opinion is so weak and vague that its probative value is outweighed by its potential prejudicial effect²⁰⁴ because fact finders are likely to overvalue such a weak and speculative opinion. My unwavering objection to the admissibility in evidence of such a “low level of confidence” opinion was reinforced by the lessons I learned from my many years of casework (not to mention the lessons handed down from my mother and her father, Professor Julius Fischhof), as well as my training in the law, all of which taught me that this type of opinion amounts to no more than a guess, and rarely an educated one given the subjectivity of forensic handwriting analysis. On those rare occasions when an objection was raised by an attorney, some courts have likewise been troubled by the inherent unreliability of an “indications opinion,” either excluding it altogether or imposing restrictions on it.²⁰⁵

In 2009, the FBI announced, to my great satisfaction, that the ASTM “indications” expression would no longer be allowed and that such levels of confidence must be reported as “no conclusion.”²⁰⁶ At the same time, the FBI laboratory announced that it would be using a five-level scale for expressing handwriting identification opinions instead of the nine-level scale contained in the ASTM E1658 standard.²⁰⁷ The FBI’s five-level scale contained a positive category permitting only two possible expressions (“identification” and “may have written”), the standard neutral category (“no

²⁰³ *Id.*

²⁰⁴ See FED. R. EVID. 403 (“The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.”).

²⁰⁵ See, e.g., *Barnett v. Am. Heritage Life Ins. Co.*, No. 06-2171, 2007 WL 7611185 (W.D. Tenn. Nov. 19, 2007); *People v. Todman*, 53 V.I. 431 (2010).

²⁰⁶ See Diana Harrison, Ted M. Burkes & Danielle P. Seiger, FBI, *Handwriting Identification: Meeting the Challenge of Science and the Law*, 11 FORENSIC SCI. COMM. (Oct. 2009), https://archives.fbi.gov/archives/about-us/lab/forensic-science-communications/fsc/oct2009/review/2009_10_review02.htm.

²⁰⁷ See DEP’T OF JUST., PROPOSED UNIFORM LANGUAGE FOR TESTIMONY AND REPORTS FOR THE FORENSIC HANDWRITING ANALYSIS DISCIPLINE 1, <https://www.justice.gov/archive/sdag/file/877746/download> (last visited Mar. 19, 2018).

conclusion”), and a negative category of two possible expressions (“elimination” and “may not have written”).²⁰⁸ Clearly, this was a significant and welcome improvement from my perspective.

To date, there is still no uniform standard for expressing handwriting identification opinions in reports and testimony. Indeed, even the FBI and the Department of Justice (DOJ) have conflicting views on this subject. The DOJ recently issued a document titled *Department of Justice Proposed Uniform Language for Testimony and Reports for the Forensic Handwriting Analysis Discipline*, which, if adopted, would “apply to Department of Justice personnel who perform examinations and/or provide expert witness testimony regarding the forensic examination of handwriting evidence.”²⁰⁹ On its face, the DOJ proposal lists a five-level scale consisting of a positive category of two opinion levels (“identification” and “qualified opinion(s) did not write”); the standard neutral category (“no conclusion”); and a negative category of two levels (“elimination” and “qualified opinion(s) did not write”).²¹⁰ However, the proposal then states that “[t]he strength or weakness of a (positive) qualified opinion will be provided by the use of any of the following terms: “Strong Probability Did Write, Probably Did Write, and Indications May Have Written,” and the corresponding negative of these three expressions may be used to express the strength or weakness of a negative qualified opinion.²¹¹ By doing so, the DOJ proposal seeks to reinstate the nine-level scale of ASTM E1658 that was specifically rejected by the FBI.

Clearly, the FDE community needs to address this critical divide by doing some serious soul-searching before deciding upon the appropriate uniform language for expressing handwriting identification opinions in reports and testimony. In my view, objective and reasonable minds will, at the very least, agree that the highly problematic and speculative “indications” opinion should not be allowed.

In the same vein, serious consideration needs to be given to proposals suggesting the avoidance altogether of the term “identification” or any other categorical conclusion regarding source attribution in favor of expressing conclusions in terms of the degree of support that exists for one proposition over another. This school of thought is predicated upon the fact that as with every other forensic pattern-matching discipline, the handwriting analyst begins with two alternative propositions (hypotheses) when comparing sets of questioned and known handwriting: (1) the two sets came from the same source (writer) or (2) the two sets came from different sources (writers). Essentially, it allows the analyst to make a subjective assessment of the

²⁰⁸ *Id.*

²⁰⁹ *Id.* at 1–2.

²¹⁰ *Id.*

²¹¹ *Id.*

strength of the evidence in support of one proposition over the other proposition based upon the analyst's observations from examining and comparing the evidence. This school of thought has found support among forensic document examiners in Australia and New Zealand,²¹² and in the 2012 Report of the NIST Expert Working Group on Human Factors in Latent Print Analysis, which recommended that latent print examiners avoid categorical conclusions about whether two sets of prints have a common source and use a more moderate expression such as, "it is far more probable that this degree of similarity would occur when comparing the latent print with the defendants' fingers than with someone else's fingers."²¹³

For me, prohibiting the "indications" opinion is essential, whereas eliminating the "identification" opinion is not, so long as it is explained to mean that the extensive degree of similarity observed in the handwriting features of the questioned and known writing warrants the belief that *the only plausible* explanation for such a concurrence is that the questioned and known writing have a common source. As I noted in Part II.C, *supra*, demonstrable evidence of shared characteristics between two sets of handwritings can provide a strong basis for a reasonable person to find that both sets of writings were made by the same person beyond a reasonable doubt.

Going forward, I firmly believe the FDE community needs to be more open to developing a standard for expressing opinions in reports and testimony that requires experts to (a) consider all the possibilities that might have accounted for the observed evidence; (b) explain why one possibility is more or less plausible than another; and (c) use qualitative probabilistic statements that more accurately convey the degree of support the observed evidence provides in favor of one proposition (possibility) over one or more alternative propositions (possibilities).

Finally, careful thought should be given to recasting the terms of the conclusions from a system of "words of estimative probability" to a system of "words of estimative surprise," consistent with and guided by Professor Risinger's insights on this issue reflected in his contribution to this symposium.²¹⁴ Professor Risinger's proposed system of estimative surprise would include a rank-order scale of words such as "mildly surprised," "surprised," "quite surprised," "greatly surprised," "astonished," "shocked," etc.²¹⁵ After reading his article, another thought came to me that might also

²¹² See Found & Bird, *supra* note 192, at 59.

²¹³ Expert Working Group on Human Factors in Latent Print Analysis. *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach*. U.S. Department of Commerce, National Institute of Standards and Technology (2012), at 134.

²¹⁴ D. Michael Risinger, *Leveraging Surprise: What Standards of Proof Imply That We Want from Jurors, and What We Should Say to Them to Get It*, 48 SETON HALL L. REV. (forthcoming 2018).

²¹⁵ *Id.* (manuscript at 19) (on file with author).

be worth pursuing, and that is to correlate each term of expression used in a scale of estimative probability with a corresponding term of expression used in a scale of estimative surprise, thereby providing a form of expression for an expert to use in “describing” or “defining” the level of confidence to be associated with the estimative probability opinion he or she has expressed in a given case. Thus, for example, if an examiner concluded that the John Doe of the known writing *did write* the questioned writing (an identification), this conclusion would have to be annotated with (accompanied by) a statement such as: “In reaching this conclusion, the examiner is saying that he would be shocked to learn that John Doe did not, in fact, write the questioned writing”; if an examiner concluded that the John Doe of the known writing *may have written* the questioned writing, this conclusion would have to be accompanied by a statement such as: “In reaching this conclusion, the examiner is saying that he would be mildly surprised to learn that John Doe did not, in fact, write the questioned writing.” Upon reflection, I believe using both forms of expression (“estimative probability” and “estimative surprise”) may provide a more transparent and reliable measure of the subjective level of confidence an expert has in his or her conclusion (opinion), and one which would be more easily understood by a lay juror, and perhaps even a judge. In any event, this issue should be explored through research studies carefully designed to evaluate how experts, judges, and lay persons (prospective jurors) interpret these various kinds of subjective expressions of expert belief (confidence).

iv. Support for and Cooperation with Properly Designed Black Box Studies

All forensic feature-comparison disciplines involve a subjective method in which perceptions and interpretations of evidence rest solely within the observer-analyst’s mind and only the conclusions derived from such perceptions and interpretations are known. Black box studies are empirical studies used to estimate the reliability of a subjective method, such as forensic handwriting analysis.

Unfortunately, the discipline of forensic handwriting analysis still lacks robust, ground truth studies that provide empirical support for the reliability of many of the tasks routinely performed by its practitioners, such as distinguishing between forged and disguised (genuine but deliberately modified) signatures. Moreover, the proficiency tests that FDEs take annually or biannually as a requirement for maintaining their individual certification or their laboratory’s accreditation status have been criticized as being too easy, not reflective of actual casework, or susceptible to test-takers colluding with one another or otherwise receiving unauthorized assistance. Hence, if the FDE community is serious about seeking to obtain trustworthy empirical data that supports the reliability of the opinions they express from performing specific handwriting examination tasks, they first need to

develop a comprehensive laundry list of all such tasks. Once this is done, empirical black box studies can be properly designed to obtain data for assessing how well FDEs perform each of those tasks. Such black box studies should be conducted by academic researchers who have no real or apparent interest in the outcome of any such black box studies.

The FDE community should unite in their efforts to support and cooperate with all academic researchers conducting such black box studies, as resistance to doing so will only provide more fodder for the discipline's critics.

If the FDE community decides to use the Modular Forensic Handwriting Method described in Part VII.A.2.b., *supra*, or a modified version of same, white box studies should also be considered, since such studies are designed to understand the factors that affect an examiner's decision at each decision point of the examination process. White box studies can be extremely useful in determining sources of error for those tasks for which black box studies produce high rates of inter-examiner variability (or even intra-examiner variability).

B. *Lawyers and Gatekeepers Need to Be Better Informed and More Diligent in Scrutinizing an Expert's Methodology*

1. The Trial Lawyer

Trial lawyers need to be more proactive and make more frequent use of *in limine* motions in seeking to exclude proffered handwriting identification opinion evidence that is unreliable. This necessarily entails having knowledge of the methodological principles and consensus standards that have been adopted by the profession for the specific handwriting task at issue in the case.

In preparing for a *Daubert* hearing or a trial, the lawyer must study the learned treatises relied upon by practitioners and all the standards adopted by the profession that are relevant to the handwriting task(s) at hand. Even with such baseline knowledge, lawyers frequently remain ill-equipped to mount an effective challenge to unreliable evidence proffered by a handwriting expert adept at obfuscating issues affecting the reliability of their opinions. Hence, lawyers need to consult with subject matter experts at the pre-trial stage and not rely solely on their cross-examination skills at a *Daubert* hearing or at trial. Presentation of contradictory expert testimony by a well-credentialed and experienced expert is always preferable and more effective in providing trial judges with the necessary baseline level of confidence to find that a proffered handwriting expert's testimony as related to the task at hand is so unfounded or illogical as to warrant finding it unreliable and inadmissible.

Reasonable access to the expertise required to expose critically flawed and unreliable proffered expert testimony is an issue of critical importance,

especially in criminal cases, where the defense customarily has inadequate financial resources to obtain such expertise, notwithstanding the Supreme Court's recognition of a constitutional right to expert assistance.²¹⁶ Trial judges may be able to remedy this inequity of resources by authorizing the payment of *reasonable* expert fees for the retention of high-quality experts by the defense, something that is generally not as great a concern for the prosecution.²¹⁷ However, this invariably requires a motion on the part of the defense attorney, optimally at the early stages of the case.

2. The Trial Judge

Judges need to take their gatekeeping role more seriously. All too often they forego their mandatory duty to screen and exclude proffered opinion evidence that is unreliable and inadmissible in favor of allowing the traditional safeguards of cross-examination and presentation of contrary evidence to expose the unreliability of an expert's opinion and trial testimony. Unfortunately, these two safeguards of our adversarial system rest upon the trial lawyer's skill level and the availability of resources to hire a well-qualified and experienced rebuttal expert, neither of which are assured. Hence, the gatekeeping role is of utmost importance in preventing unreliable opinion evidence from reaching the fact finder.

The cause for concern in relying upon gatekeepers to exclude unreliable opinion evidence is that most trial judges are not technically proficient when it comes to most forensic science disciplines. They have generally not been exposed to, or trained to identify, the kinds of flaws in examining or evaluating evidence that produce unreliable and exaggerated expert opinions, even in the case of pattern-matching disciplines such as handwriting analysis. Unless judges appreciate and comprehend the reliability issues involved in challenges to proffered opinion evidence, it is unlikely that those issues will be addressed in a meaningful way.

To facilitate and enhance the performance of their gatekeeping responsibility, as outlined in Part VI, *supra*, judges should consider implementing the following practical measures to assist the court in assessing the reliability of proffered opinion evidence:

²¹⁶ *Ake v. Oklahoma*, 470 U.S. 68, 82–87 (1985). See also Paul C. Giannelli, *Ake v. Oklahoma: The Right to Expert Assistance in a Post-Daubert World*, 89 CORNELL L. REV. 1305 (2003).

²¹⁷ See James Fanelli, *Manhattan DA's Office Paid Unbelievable Amount to Psychiatrist Who Assessed Convicted Etan Patz Killer*, N.Y. DAILY NEWS (July 29, 2017), <http://www.nydailynews.com/new-york/prosecutors-paid-massive-amount-expert-etan-patz-case-article-1.3367894>. The New York Daily News reported, having learned through a Freedom of Information Law request that “the Manhattan District Attorney’s Office paid a forensic psychiatrist \$536,940 for his work as an expert witness in the two trials of Etan Patz’s killer” whereas “[the defendant’s] court-appointed attorneys and paralegals were paid a total of \$803,000 according to the city Finance Department.” *Id.*

2018]

A DOCUMENT EXAMINER TELLS ALL

717

- a) routinely issue pre-trial orders directing all experts to submit written reports or affidavits detailing the reasoning and methods underlying their opinions; and
- b) appoint a court-appointed subject matter expert in those instances when the specialized knowledge of a nonpartisan expert would assist the court in properly assessing the reliability of proffered opinion evidence.

CONCLUSION

Has there been progress since the 2009 NAS Report? Yes, but not enough, despite the NIST-OSAC efforts to improve the various disciplines of forensic science, especially those involved in the examination and comparison of features associated with physical or trace evidence. The FDE community has engaged in a lot of talk, but far too little action, and the problems associated with the reliability of handwriting identification opinion evidence, as described and illustrated in this Article, still prevail.

Hopefully the information provided in the first five parts of this Article will encourage and ignite members of the FDE community, as well as members of the bench and the bar, to become more proactive in addressing and resolving the critical issues that continue to adversely affect the reliability of handwriting identification opinion evidence by, at the very least, implementing the changes recommended in Part VII of this Article.