

2013

Intuition Versus Algorithm: The Case of Forensic Authorship Attribution

Lawrence M. Solan, Ph.D., J.D.

Follow this and additional works at: <https://brooklynworks.brooklaw.edu/jlp>

Recommended Citation

Lawrence M. Solan, Ph.D., J.D., *Intuition Versus Algorithm: The Case of Forensic Authorship Attribution*, 21 J. L. & Pol'y (2013).

Available at: <https://brooklynworks.brooklaw.edu/jlp/vol21/iss2/13>

This Article is brought to you for free and open access by the Law Journals at BrooklynWorks. It has been accepted for inclusion in Journal of Law and Policy by an authorized editor of BrooklynWorks.

INTUITION VERSUS ALGORITHM: THE CASE OF FORENSIC AUTHORSHIP ATTRIBUTION

*Lawrence M. Solan**

I. INTRODUCTION

On November 6, 2012, Barack Obama was reelected President of the United States, having defeated his Republican opponent, Mitt Romney. The vote in the Electoral College—the official body that votes on a state-by-state basis—was decisive: 332–206.¹ Obama also won the national popular vote by a margin of about 4,850,000 votes (50.9% to 47.1%).² But Obama’s winning by a comfortable margin is not what many of the pundits on television were predicting. Some announced that Romney would win the election, including predictions that he

* Don Forchelli Professor of Law, and Director, Center for the Study of Law, Language and Cognition, Brooklyn Law School. My thanks to Silvia Dahmen, Lorna Fadden, Roger Shuy and Ben Zimmer for valuable comments on earlier drafts, and to the participants in the Authorship Attribution Workshop held at Brooklyn Law School in October 2012, sponsored by the National Science Foundation. This article is a substantially expanded and revised version of my commentary, *Ethics and Method in Forensic Linguistics*, PROC. INT’L ASS’N FORENSIC LINGUISTS’ TENTH BIENNIAL CONF., 2012, at 362, available at <http://www.forensiclinguistics.net/iafl-10-proceedings.pdf>. Research on this article was sponsored by a Dean’s Summer Research Stipend from Brooklyn Law School.

¹ *2012 Electoral College Results*, NARA, <http://www.archives.gov/federal-register/electoral-college/2012/election-results.html> (last visited Apr. 6, 2013).

² *2012 Presidential Election*, NARA, <http://www.archives.gov/federal-register/electoral-college/2012/popular-vote.html> (last visited Apr. 6, 2013).

would win by a landslide.³ Most guessed that the election would be much closer than it turned out to be.

If the television pundits were all over the lot and mostly wrong, those who used sophisticated computational techniques to draw inferences from polls fared much better. An article in the *New York Times* shortly after the election put it this way:

It was not on any ballot, but one of the biggest election contests this week pitted pundits against pollsters. It was a pitched battle between two self-assured rivals: those who relied on an unscientific mixture of experience, anecdotal details and “Spidey sense,” and those who stuck to cold, hard numbers.

When the results were tabulated, it became clear that data had bested divination.⁴

Perhaps most prominent among the pollsters was *New York Times* blogger Nate Silver. As of the morning of the election, his “FiveThirtyEight” blog predicted that Obama would receive 313 electoral votes to Romney’s 225, and that Obama had a 90.9% chance of winning the election.⁵ Silver also predicted that Obama would win the popular vote by 2.5 percentage points.⁶ He underestimated the margin of victory slightly in each measure, but not by much, and did dramatically better than did the pundits. Silver’s success made him a celebrity of sorts, including an appearance on *The Daily Show with Jon Stewart*.⁷

³ See Benny Johnson, *Romney Landslide: Here Are the Biggest Names Predicting It*, THEBLAZE.COM (Nov. 4, 2012, 3:37 PM), <http://www.theblaze.com/stories/romney-landslide-here-are-the-biggest-names-predicting-it-how-it-will-happen/>. Among such predictors were Dick Morris, Karl Rove, Larry Kudlow, Joe Scarborough, and George Will. For quotes from these pundits, see *id.*

⁴ Michael Cooper, *Election Result Proves a Victory for Pollsters and Other Data Devotees*, N.Y. TIMES, Nov. 8, 2012, at P8.

⁵ Nate Silver, *FiveThirtyEight: When Internal Polls Misperform, a Whole Campaign May Be to Blame*, N.Y. TIMES (Dec. 1, 2012), <http://fivethirtyeight.blogs.nytimes.com/2012/12/01/when-internal-polls-misperform-a-whole-campaign-may-be-to-blame/>.

⁶ *Id.*

⁷ *The Daily Show with Jon Stewart* (Comedy Central television broadcast Nov. 7, 2012), available at <http://www.thedailyshow.com/watch/wed->

His book *The Signal and the Noise* is a tribute to the triumph of algorithm over intuition.⁸

It would be a mistake, however, to dismiss intuitive expert judgment generally, just because it fails at predicting the results of presidential elections. Not all expert opinion based upon experience can be reduced to “divination.” Through repeated experience, people develop expertise of all kinds, ranging from chess playing⁹ to medical diagnosis.¹⁰ No one accuses the best of such people of using a divining rod simply because they do not rely on computer algorithms. Moreover, we make judgments all the time about what is likely to happen next, including, for example, the judgment that it is safe to cross the street when the traffic signal favors us and the cars are all stopped. Most of the time, there is no computer algorithm with which we can compare our rate of success, but we have a good sense—confirmed by repeated experience—that we are making the right decision.

On the other hand, the use of algorithms seems to neutralize some obvious biasing factors that plague the pundits routinely. Why is it that experts paid by Fox News (a Republican-oriented network) predicted a Romney victory, whereas those paid by MSNBC (a Democrat-oriented network) predicted that Obama would win? One possibility is that most of the pundits are sufficiently corrupt to misstate their actual views if they are paid enough to do so. More likely, though, their prior commitments contribute to what information they regard as significant and color their analyses, which are sincere. This phenomenon, called confirmation bias, is well studied by psychologists. It is an

november-7-2012/nate-silver.

⁸ See generally NATE SILVER, *THE SIGNAL AND THE NOISE* (2012) (investigating how statisticians distinguish meaningful indicators in ever-increasing amounts of data in order to make accurate predictions).

⁹ Fernand Gobet & Neil Charness, *Expertise in Chess*, in *THE CAMBRIDGE HANDBOOK OF EXPERTISE AND EXPERT PERFORMANCE* 523, 532–34 (K. Anders Ericsson et al. eds., 2006).

¹⁰ Geoff Norman et al., *Expertise in Medicine and Surgery*, in *THE CAMBRIDGE HANDBOOK OF EXPERTISE AND EXPERT PERFORMANCE*, *supra* note 9, at 339, 350.

“unwitting selectivity in the acquisition and use of evidence.”¹¹ The networks engage just those experts whose views are most likely to reinforce the views of their audience. Similarly, why is it that the television networks so routinely predicted a close election? Could it be relevant that these privately-owned media outlets make their living selling audiences to advertisers, and it is in their interest to maintain electoral drama for as long as is feasible? No doubt confirmation bias plays a role here as well.

With the election in mind, let us move to forensic authorship attribution. In his essay on the current state of the field, Professor Ronald Butters reminds us, with insight and candor, that forensic linguists, like practitioners in most areas of forensic science, have done more to advance their field substantively than they have done to advance it ethically.¹² The program he suggests is an ambitious one. Butters complains that forensic authorship attribution lacks not only a set of agreed understandings about methodology but also lacks, and is in need of, standards¹³ sufficient to ensure the exclusion of bogus conclusions based on inadequate data.¹⁴ In this regard, Butters places methodology beyond mere practice and elevates it to the realm of the ethical: it is simply wrong for a profession to go about its business without some verification that it is doing a good job. Professor Joseph Sanders raises similar points in an essay on the ethical duties of expert witnesses more generally.¹⁵ What could be more important than making sure that those academics, whose “day jobs” are to seek the truth, do more

¹¹ Raymond S. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 2 REV. GEN. PSYCHOL. 175, 175 (1998).

¹² Ronald R. Butters, *Retiring President's Closing Address: Ethics, Best Practices, and Standards*, PROC. INT'L ASS'N FORENSIC LINGUISTS' TENTH BIENNIAL CONF., 2012, at 351–52. The essay is the text of the Presidential Address delivered by Professor Butters at the meetings of the International Association of Forensic Linguists, Aston University, Birmingham, U.K., July 2011.

¹³ *Id.* at 352–53, 356.

¹⁴ *See id.* at 356.

¹⁵ Joseph Sanders, *Expert Witness Ethics*, 76 FORDHAM L. REV. 1539, 1583 (2007) (calling for codes of ethics for individual fields to guide experts as to their responsibility in taking an appropriate epistemological stand toward their testimony).

good than harm when they enter the courtroom with the express task of presenting analysis that will affect the lives of others in profound ways?

This essay explores some of the issues that Butters raises in the context of forensic authorship attribution analysis and that others have raised for some time in the context of other forensic sciences that rely on trace evidence.¹⁶ My first point is that the conflict of interest inherent in expert forensic testimony—especially by those who make their livings, or at least a significant part of their livings, as consulting experts—can indeed best be remedied by the development of methods that are demonstrably both diagnostic and replicable. For those who rely upon judgments of coauthorship based on their knowledge of linguistic features and upon a sense that a large cluster of differences or commonalities in a particular case cannot be a matter of accident, research into methodology should be a top priority. Proficiency testing may take the place of the development of replicable methods in the short run, but the best direction for the field is to demonstrate that methods work and are not highly dependent on the skill of the practitioner alone.

My second point is that work in computer science and computational linguistics is moving toward answering many of the specific questions that Butters raises about particular standards in the field. Such matters as how much data are needed for valid conclusions to be drawn are commonplace in statistics and modeling, and can easily enough be transported to forensic linguistic application. I end this essay with some brief conjecture about why the field does not appear to have moved ahead quickly with respect to some of these questions and what it might do to adjust its course.

II. LUCY AND LACY: TWO STYLES OF EXPERT ANALYSIS

Those who engage regularly in expert consultation, and especially in expert testimony, have an inherent conflict of

¹⁶ See, e.g., 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, 27–42 (2002).

interest. It is sometimes referred to as the “hired gun” syndrome, and it stems from the fact that testifying experts are encouraged to render opinions useful to the party that hires them and are subject to confirmation bias in any event. Consider the following vignette about expert witness Lucy:

Lucy is a professor of computational linguistics and currently has a grant-funded project on authorship identification, which she hopes will have practical application eventually. Last month, a lawyer phoned Lucy, saying he had heard of her work, and asked her if she would be willing to apply it to a legal case and possibly testify as an expert. Lucy was intrigued. She took the case, analyzed it according to the methods that she had developed, and concluded, by virtue of applying her algorithm, that the questioned document the lawyer presented was very unlikely to have been written by the person to whose known writings she had compared it. In her lab, Lucy was correct 88% of the time when she conducted this kind of analysis this way. She told the lawyer that she would be happy to testify to all of this, as she continues to work in her lab to improve the 88% rate of accurate rejection of authorship.

Now compare Lucy to Lacy:

Lacy is a forensic linguistic consultant. From time to time she takes authorship attribution cases. Lacy does not conduct her work computationally. Rather, she has a set of thirty-six stylistic markers by which she analyzes all documents that come to her. She has found from past experience that when the documents are long enough for comparison, some of these thirty-six markers will tend either to co-occur between a questioned document and a reference set or be noticeably different between them. There is sometimes controversy about whether her testimony will be permitted, but when she is allowed to testify, her testimony is generally convincing.

At first glance, we might prefer Lucy. After all, we know how good her methods are, making it less likely that she is a hired gun. With Lacy, in contrast, we must rely on her

persuasive rhetoric and the intuitive appeal of the data she presents.

Herein lies the problem: we have no idea which expert does a better job. It may well be that Lacy limits herself to the kinds of problems that she is certain to get right and that her success rate exceeds Lucy's not unimpressive 88%. On the other hand, it may be that Lacy gets a lot of slack from her charisma and the intuitive appeal of her analyses and that her success rate is far lower than Lucy's.

Over the past two decades, forensic linguistics, I believe, has developed as a field with more Lacys than Lucys, and this has led to some of the problems that Butters observes. Many involved in the field—especially authorship attribution specialists who rely on stylistic markers—conduct little or no laboratory work. This is true both of independent consultants and of academics who self-identify as forensic linguists. The result is a dearth of serious research, provoking reasonable questions about the legitimacy of the conclusions reached. As noted below, proficiency testing may be at least a partial solution to this problem, but no such testing currently takes place. At the same time, somewhat disconnectedly, computer scientists and computational linguists have been developing algorithms that more and more successfully predict authorship, but much of this has not yet made its way to forensic application.¹⁷

This tension was not always so pronounced. The history of “voiceprint” analysis provides quite a different story. During the 1960s, an employee of Bell Labs, which invented the sound

¹⁷ For the state of current research, see Shlomo Argamon & Moshe Koppel, *A Systemic Functional Approach to Automated Authorship Analysis*, 21 J.L. & POL'Y 299 (2013); Moshe Koppel et al., *Authorship Attribution: What's Easy and What's Hard?*, 21 J.L. & POL'Y 317 (2013); Efstathios Stamatatos, *On the Robustness of Authorship Attribution Based on Character N-Gram Features*, 21 J.L. & POL'Y 421 (2013). For an overview, see Carole E. Chaski, *Author Identification in the Forensic Setting*, in THE OXFORD HANDBOOK OF LANGUAGE AND LAW 489 (Peter M. Tiersma & Lawrence M. Solan eds., 2012); Patrick Juola, *Authorship Attribution*, 1 FOUND. & TRENDS IN INFO. RETRIEVAL 233 (2008). For general discussion, see Lawrence M. Solan, *The Expert Witness Meets the Adversarial System*, in THE ROUTLEDGE HANDBOOK OF FORENSIC LINGUISTICS 395 (Malcolm Coulthard & Alison Johnson eds., 2010).

spectrograph, began to make extravagant claims about the ability of such devices to distinguish one voice from another, much the way fingerprints were (and still largely are) seen as distinguishable from one person to the next.¹⁸ Police laboratories received training in the use of the new technology, about which there was considerable excitement.¹⁹ Then, in 1979, the National Research Council issued a devastating report, pointing out that there had not been adequate testing to determine how well spectrography can be used to distinguish one voice from the other in forensic settings.²⁰ The report noted:

The degree of accuracy, and the corresponding error rates, of aural-visual voice identification vary widely from case to case, depending upon several conditions including the properties of the voices involved, the conditions under which the voice samples were made, the characteristics of the equipment used, the skill of the examiner making the judgments, and the examiner's knowledge about the case. Estimates of error rates now available pertain to only a few of the many combinations of conditions in real-life situations. These estimates do not constitute a generally adequate basis for a judicial or legislative body to use in making judgments concerning the reliability and acceptability of aural-visual voice identification in forensic applications.²¹

The leaders in the effort to make sure that linguistic science, if used in law enforcement efforts, would meet the high standards of science itself were chiefly academic linguists, with

¹⁸ See Lawrence G. Kersta, *Voiceprint Identification*, 196 NATURE 1253, 1257 (1962). For discussion of this history, see LAWRENCE M. SOLAN & PETER M. TIERSMA, SPEAKING OF CRIME: THE LANGUAGE OF CRIMINAL JUSTICE 140, 140–46 (2005) and Lawrence Solan & Peter Tiersma, *Hearing Voices: Speaker Identification in Court*, 54 HASTINGS L.J. 373, 416–26 (2003).

¹⁹ See SOLAN & TIERSMA, *supra* note 18, at 140.

²⁰ NAT'L RESEARCH COUNCIL OF THE NAT'L ACADS., ON THE THEORY AND PRACTICE OF VOICE IDENTIFICATION 58 (1979).

²¹ *Id.* at 60.

special efforts by the distinguished phonetician, Peter Ladefoged.²²

Fast-forward to 2009, when the National Research Council came out with another devastating report, this time concerning forensic identification science in the United States more generally. The report decried the absence of scientific integrity in forensic identification procedures, much as the earlier report did with respect to speaker identification technology used at the time:

A body of research is required to establish the limits and measures of performance and to address the impact of sources of variability and potential bias. Such research is sorely needed, but it seems to be lacking in most of the forensic disciplines that rely on subjective assessments of matching characteristics. These disciplines need to develop rigorous protocols to guide these subjective interpretations and pursue equally rigorous research and evaluation programs.²³

Yet now, it is not the linguistic academic community taking the lead in remedying this situation on behalf of linguists who do not want to see the legal system making excessive claims about the forensic application of the language sciences. Rather, it is chiefly legal academics with expertise in the area of scientific evidence taking the lead, with the focus not on linguistics in particular but on the forensic identification sciences generally.²⁴ What has happened between 1979 and 2009? In 1979, there was no field of forensic linguistics, or at least not much of one. Linguists were occasionally called to testify as experts in court, but they did so because their academic expertise serendipitously crossed paths with a legal issue, much the way an academic physicist or engineer might be called upon to provide expert testimony. Linguists were in the business of being linguists, and

²² See SOLAN & TIERSMA, *supra* note 18, at 140–41.

²³ NAT'L RESEARCH COUNCIL OF THE NAT'L ACADS., STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 8 (2009).

²⁴ See, e.g., Jonathan J. Koehler, *If the Shoe Fits They Might Acquit: The Value of Forensic Science Testimony*, 8 J. EMPIRICAL LEGAL STUD. 21 (2011); Risinger et al., *supra* note 16; Sanders, *supra* note 15.

the legal system was as much a novelty to them as they were a novelty to the legal system. As late as 1994, Judith Levi, one of the founders of the field of forensic linguistics, wrote the inaugural article in the then-brand-new journal *Forensic Linguistics* (now, *The International Journal of Speech, Language and the Law*), explaining to the linguistics community at large the growing acceptance of linguistic experts in court.²⁵

Since that time, things have developed considerably. With the growth of undergraduate and graduate programs in forensic linguistics,²⁶ many academics devote much of their time to applying linguistic knowledge in legal settings, as do consulting linguists without academic affiliations. The gap between the academic community that once policed its field for abuse by the legal system and practitioners within the legal system has shrunk considerably.

Some in the language and law community have recognized this gap between theory and practice to be a healthy one and have attempted to maintain it in their description of how linguists engaged in courtroom testimony should view their work. Peter Tiersma, a law professor and linguist (with whom I frequently write), suggests that the field is at its best when the reluctant scholar is asked to share her expertise with the court for the benefit of reaching a proper result in a dispute in which the expert has no interest, either intellectual or otherwise.²⁷ Roger Shuy, a distinguished scholar in applied linguistics who has been a prolific writer in the field, states the forensic linguist's ideal role more fully:

[*Forensic linguistics*] has become a useful way to refer to the use of linguistics knowledge where there are data that

²⁵ See Judith N. Levi, *Language as Evidence: The Linguist as Expert Witness in North American Courts*, 1 FORENSIC LINGUISTICS 1 (1994). For subsequent history, see Peter M. Tiersma & Lawrence M. Solan, *The Linguist on the Witness Stand: Forensic Linguistics in American Courts*, 78 LANGUAGE 221 (2002).

²⁶ Such programs exist at Cardiff University, Aston University (Birmingham, UK), Universitat Pompeu Fabra (Barcelona), and Hofstra University.

²⁷ Peter M. Tiersma, *Linguistic Issues in the Law*, 69 LANGUAGE 113, 122 (1993).

serve as evidence. But I have some concerns about the term itself, because it seems that when one does “forensic linguistics” one is simply doing linguistics, a type of applied linguistics, in fact.²⁸

These views are attractive. To the extent that they are descriptively accurate views of the field, they explain how it was that the academic phoneticians were the ones who shut down the voiceprint craze of the 1960s and ‘70s.

Now, however, expert testimony on questions of language goes beyond ordinary research in linguistics into areas developed by those interested in forensic linguistics as its own discipline.²⁹ Individuals, sometimes without a great deal of training in linguistics, generally become skilled in areas of language analysis developed particularly for consumption by the legal community.³⁰ For reasons stated earlier, many practitioners have little motivation to police their own methodologies along the conventional scientific criteria of validity and reliability and typically do not engage in analysis of methods used by others, apart from disagreement within a particular case in which they have taken opposing positions. The result of this development is that it is not always possible to judge forensic testimony against ordinary practices among linguists, because linguists do not ordinarily engage in the activities that generate the expert testimony.

Does this amount to an ethical issue? It probably does if neither methodological testing nor proven individual proficiency forms a sufficient scientific basis to accept some of what passes for linguistic expertise. Moreover, to the extent that this lack of foundation results from the dearth of research that is itself a product of this conflict, then it is the fruit of a conflict of interest and is an ethical issue for this reason. Law professor and philosopher Susan Haack puts it this way:

²⁸ ROGER W. SHUY, *LINGUISTICS IN THE COURTROOM: A PRACTICAL GUIDE* 3 (2006).

²⁹ See Ronald R. Butters, *The Forensic Linguist's Professional Credentials*, 16 *INT'L J. SPEECH LANGUAGE & L.* 237 (2009).

³⁰ *Contra* SHUY, *supra* note 28, at 3.

Distinguishing genuine inquiry, the real thing, from pseudo-inquiry or “sham reasoning,” C.S. Peirce—a working scientist as well as the greatest of American philosophers—wrote that “the spirit . . . is the most essential thing—the motive”; that genuine inquiry consists in “actually drawing the bow upon truth with intentness in the eye, with energy in the arm.” For the same reason, I am tempted to write of advocacy “research” (in scare quotes); for it is something of a stretch to call advocacy research “research” at all. Advocacy “research” is like inquiry insofar as it involves seeking out evidence. But it is part of an advocacy project insofar as it involves seeking out evidence favoring a predetermined conclusion; and it is undertaken in the spirit, from the motive, of an advocate. In short, it is a kind of pseudo-inquiry.³¹

At the same time, as noted previously, there is no reason to conclude that intuitive expertise based on experience and insight fares any better or worse than does algorithmic expertise. Intuitive expertise is not necessarily unreliable. On the contrary, it is clear that at least in some settings, people are able to form sophisticated mental models of situations about which they are experts and to weigh relevant factors with great accuracy, notwithstanding that they are unable to describe how they did it. For example, Malcolm Gladwell, in his book *Blink*, made famous the story of the J. Paul Getty Museum’s acquisition of a 2,000-year-old Greek sculpture—a kouros—which is a rare thing to acquire.³² The museum did its due diligence carefully, investigating the sculpture’s provenance over the centuries, engaging experts to examine the marble with microscopes, and so on.³³ But the day of reckoning came when the museum’s curator began inviting various experts in classical sculpture, none of whom felt that the sculpture was authentic, and one of whom remarked that seeing it caused in him a wave of “intuitive

³¹ Susan Haack, *What’s Wrong with Litigation-Driven Science? An Essay in Legal Epistemology*, 38 SETON HALL L. REV. 1053, 1071 (2008).

³² MALCOLM GLADWELL, *BLINK* 3 (2005).

³³ *Id.* at 3–4.

repulsion.”³⁴ For the most part, the experts could not put into words exactly what was bothering them, but deep within themselves, they knew.³⁵ Later, all agreed that the kouros was a kouros copy.³⁶

The kouros story is particularly relevant here because it is, in a sense, a matter of authorship attribution: who sculpted the kouros that the museum had bought? Thus, we may wish to ask about the nature of the knowledge that the experts brought with them to the task that led to their negative reactions. As Daniel Kahneman points out,³⁷ celebrations of gut-reaction decision making, such as Malcolm Gladwell’s *Blink*, also contain stories of terrible diagnostic failure, including the misdiagnosis of heart attacks.³⁸ And we’ve already seen how well political experts do at predicting election results. Many in the business of predicting the future of the economy are on a par with the political pundits. It thus appears that intuitive expertise is neither all good nor all bad as a diagnostic tool.³⁹

Psychologists have devoted a great deal of study to the question of expert intuition, in areas as diverse as the thinking of chess masters, medical diagnosis and treatment by physicians, neonatal intensive care nursing, and decision making about firefighting.⁴⁰ What appears to separate the intuitive experts from the rest of us is a deep base of knowledge that has enabled them to build mental models⁴¹ that are so robust that they can be both accessed and expanded with ease. Chess masters do far better

³⁴ *Id.* at 5–6.

³⁵ *Id.*

³⁶ *Id.* at 7.

³⁷ DANIEL KAHNEMAN, THINKING, FAST AND SLOW 235–36 (2011).

³⁸ GLADWELL, *supra* note 32, at 130.

³⁹ For a discussion of how scientific analysis and subjective analysis blend in diagnosis in many domains, see John A. Swets et al., *Psychological Science Can Improve Diagnostic Decisions*, 1 PSYCHOL. SCI. IN PUB. INTEREST 1 (2000).

⁴⁰ For discussion of the circumstances in which intuitive expertise is most likely to prevail, see Daniel Kahneman & Gary Klein, *Conditions of Intuitive Expertise: A Failure to Disagree*, 64 AM. PSYCHOLOGIST 515, 522 (2009).

⁴¹ For an overview of what constitutes a mental model, see P.N. JOHNSON-LAIRD, MENTAL MODELS 10–12 (1983).

than do novices in recalling the pieces in familiar chess configurations but generally do not do significantly better in recalling the pieces placed at random on a chess board.⁴² Moreover, the more such configurations are familiar to them—because they devote their lives to studying these configurations—the broader the range of improved recall.⁴³ By the same token, experienced doctors, when asked the basis of a diagnosis, do not resort to basic science, even though they were taught and know the mechanisms underlying disease. In a hard case they may resort to their basic training, but as a general matter, they have developed, based on huge numbers of observations, a wealth of knowledge of what combinations of symptoms and conditions are indicative of disease and can do this with great accuracy, even when they are not able to articulate how it is that they assigned particular weights to particular symptoms in a particular setting.⁴⁴ And the same holds true for the impressive record of neonatal intensive care nurses, fire commanders, and other such experts.⁴⁵ Each group of experts develops mental models capable of distinguishing the successful from the unsuccessful based upon repeated exposures to a wealth of scenarios.

Thus, while Lacy relies upon intuitive judgment more than does Lucy, depending upon how conservative Lacy is in her willingness to offer expert opinions, her rate of error may be lower than that of Lucy. Yet, when given a choice, the legal system typically opts for Lucy-like expertise, not because we know that she is more likely to be correct, but because we at least know how likely she is to be right, reducing the probability

⁴² Gobet & Charness, *supra* note 9, at 526–27.

⁴³ *Id.* at 526. Nonetheless, the machines have caught up to and passed the chess masters. For an interesting account, see SILVER, *supra* note 8, at 262–93 (describing Garry Kasparov's games with computer chess programs during the 1980s and 90s and his eventual loss to Deep Blue in 1997).

⁴⁴ Norman et al., *supra* note 10, at 346.

⁴⁵ For a description of some of this research and a theory of what makes intuitive expertise successful in these circumstances, see Karol G. Ross et al., *Professional Judgments and "Naturalistic Decision Making,"* in THE CAMBRIDGE HANDBOOK OF EXPERTISE AND EXPERT PERFORMANCE, *supra* note 9, at 403, 403–15.

that cognitive biases or a witness's compelling personality will play too great a role in the outcome of a case.⁴⁶ Making the case for algorithmic expertise more compelling, people are much better at recognizing the biases of others than they are at recognizing their own biases.⁴⁷ Thus, encouraging experts to recognize and stave off the temptation of becoming too much a team player is not likely to be an adequate solution to the problem of bias.

The literature on the nature of intuitive expertise raises another concern with respect to authorship attribution. Expert opinion testimony is admissible only if the expert's scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue.⁴⁸ It is not clear how much of the expert opinion of intuitive experts on authorship attribution is a matter of expertise. To see why this is the case, let us look at a Lacy-like analysis taken from an article by Australian linguist Robert Eagleson.⁴⁹ Eagleson describes a case in which a woman supposedly left a farewell letter to her husband, typed on the family typewriter, when she ran off with another man.⁵⁰ The police believed, however, that the husband had written the letter, and had done away with the wife.⁵¹ A linguist was called in to compare the style of the farewell letter with the style of writing of documents known to be written by the husband and documents known to be written by the wife.⁵²

⁴⁶ See Joseph Sanders, *Kumho and How We Know*, 64 LAW & CONTEMP. PROBS. 373, 374–75, 393 (2001). The propensity to overstate the role of character and to understate the circumstances in which an individual acts is called “the correspondence bias” in the psychological literature. See Daniel T. Gilbert & Patrick S. Malone, *The Correspondence Bias*, 117 PSYCHOL. BULL. 21 (1995); see also SOLAN & TIERSMA, *supra* note 18, at 29–32 (discussing evidentiary standards in the *Daubert* age).

⁴⁷ Emily Pronin et al., *The Bias Blind Spot: Perception of Bias in Self and Others*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 369, 369–81 (2002).

⁴⁸ FED. R. EVID. 702.

⁴⁹ Robert Eagleson, *Forensic Analysis of Personal Written Texts: A Case Study*, in LANGUAGE AND THE LAW 362 (John Gibbons ed., 1994).

⁵⁰ *Id.*

⁵¹ *Id.*

⁵² *Id.*

The results were dramatic. The husband's known writings and the farewell letter shared a number of nonstandard spellings (individual words, capitalization of common nouns, lower case proper nouns, use of apostrophe), syntactic structures, word structures, and punctuation nuances.⁵³ The wife's writings had none of these features.⁵⁴ To take one example, both the husband's writings and the farewell letter used the present tense when it would have been appropriate to use the past tense ("He threaten me.").⁵⁵ After the results of this analysis were presented to the husband, he confessed.⁵⁶

But there is a problem here. The grouping of similarities and differences indeed requires some sophistication in the analysis of language. It is unlikely that someone not trained in linguistics would have come up with this array. Once the linguistic expert did so, however, there was no particular science behind the inference that the husband was more likely than the wife to have written the farewell letter. It only makes sense given the array of similarities with the husband's style and differences from the wife's, but it makes sense because of what our common sense notions tell us about how likely people are to be consistent about such aspects of their writing. Missing is the kind of experience that the doctors, neonatal nurses, chess players, and others describe in which the similar patterns are presented to them over and over again with the results known quickly. Other Lacy-like examples show the same characteristics—a substantial, and often intuitively convincing, number of similarities between a questioned document and the writings of a suspect, with no serious science underlying the inference of authorship identification.⁵⁷

This leaves the legal system with three choices: it can accept the expert testimony, opinion and all; it can reject the expert testimony on similarities and differences entirely; or it can admit

⁵³ *Id.* at 371–72.

⁵⁴ *Id.*

⁵⁵ *Id.* at 367–68.

⁵⁶ *Id.* at 373.

⁵⁷ See, for example, the discussion of the Unabomber case in SOLAN & TIERSMA, *supra* note 18, at 159–64.

the testimony, but permit the jury to draw the inferences of identity for which there is not an adequate scientific basis. Each of these options has its own difficulties.

The problem with the first option is that there is no basis for crediting the opinion of the expert. In the next section, I recommend proficiency testing to legitimize expert testimony in these circumstances.

The second option, while seemingly attractive, especially to those who favor an algorithmic approach, also comes with a price. It takes from the jury the ability to evaluate evidence with more sophistication. Take, for example, a point raised by Malcolm Coulthard⁵⁸: the government offers evidence in a fraud case that a letter whose authorship is in question contains the same kinds of linguistic nuances (spelling errors, punctuation, and so on) as do the known writings of the defendant.⁵⁹ No expert is needed.⁶⁰ Now, the defense offers an expert to demonstrate that these same features are present in the writings of other possible authors.⁶¹ Only the expert conducted the evaluation of the documents, so without his testimony, they will not be before the jury.⁶² It would seem that justice is better served if the expert is permitted to point out similarities with other potential authors, regardless of there being no algorithm that can demonstrate a likelihood of actual authorship. After all, that is exactly what the prosecution has done without an expert in its efforts to implicate the defendant.

The third option—permitting the expert to point out similarities and differences, but not to issue an opinion on attribution—also has its problems. The approach has initial appeal. On the one hand, the comparison seems to be relevant to the outcome. On the other hand, we have no way of knowing how good the expert's lay intuitions are likely to be. No studies have been conducted to tell us. They may be right most of the

⁵⁸ Malcolm Coulthard, *On Admissible Linguistic Evidence*, 21 J.L. & POL'Y 441 (2013).

⁵⁹ *Id.*

⁶⁰ *Id.*

⁶¹ *Id.*

⁶² *Id.*

time when so many features are either shared or differentiated, but we simply do not know. In a different context, I recommended that courts accept a “tour guide” approach to expert linguists testifying about meaning.⁶³ When each side appears to have proposed a reasonable interpretation of legally relevant language, a linguist may point out the various plausible interpretations and explain how they derive from ordinary linguistic processes.⁶⁴ This kind of testimony poses little danger, since the judge and jury are perfectly capable, based on their intuitions as speakers of English, to determine whether the linguist’s testimony accurately reflects their own judgments about the range of possible meanings.⁶⁵ Moreover, once the range of plausible interpretations is brought out, the linguist’s expert opinion about meaning is largely superfluous, since the expert will have put the jury on an equal footing with him by virtue of the testimony.⁶⁶

Authorship attribution is different, however. The goal of the expert is not to make jurors sensitive to the full range of their intuitions about authorship but rather to determine who wrote the questioned document. We do not know, however, to what extent the expert testimony on similarities and differences is helpful and how much it leads jurors to intuitive judgment without adequate basis to determine whether the similarities and differences that appear so telling have any real predictive force.

Thus, as Edward Cheng points out, “[t]he heavy-lifting in developing an authorship attribution technique should occur in the lab, long before it is applied in a legal case.”⁶⁷ While that is happening, however, courts are faced with the uncomfortable dilemmas described above. Let us now turn to how the American legal system has reacted to these issues and how the field might develop to increase its efficacy in court.

⁶³ Lawrence M. Solan, *Linguistic Experts as Semantic Tour Guides*, 5 FORENSIC LINGUISTICS 87 (1998).

⁶⁴ *Id.* at 94–95.

⁶⁵ *Id.* at 95.

⁶⁶ *Id.* at 92.

⁶⁷ Edward K. Cheng, *Being Pragmatic About Forensic Linguistics*, 21 J.L. & POL’Y 541, 550 (2013).

III. JUDICIAL REACTIONS TO AUTHORSHIP ATTRIBUTION EVIDENCE

This section comments on what the American courts have been saying about authorship identification expertise in the United States. However, as Peter Tiersma and I have pointed out,⁶⁸ relying upon published opinions to draw conclusions about evidentiary decisions by American courts is not likely to present a fair sense of what actually happens in trial courts. The case law provides a very deferential approach to appeals of decisions on the admissibility of expert opinion evidence,⁶⁹ and most (but not all) of the appeals will occur in the context of the losing party having been denied in their application to have an expert testify. The result is that most evidentiary decisions in published opinions by appellate courts are affirmances of the decision of the trial court to exclude an expert. An academic, or for that matter, a lawyer or judge, who relies on these opinions will not have any idea in how many cases experts have indeed testified at trial. Yet such testimony will occur when both sides call experts on the same issue, when one side calls an expert without objection from the other side, or when the offer of an expert survives a motion to exclude, but the case does not result in a published opinion, at least not on that issue.

Prominent examples have appeared in the press in 2011 and 2012. For example, Robert Leonard and Ronald Butters each testified in the Chicago murder case, *People v. Coleman*.⁷⁰ Coleman was accused of killing his wife.⁷¹ Part of the prosecution's story was that he had written various threat letters, one of them painted as graffiti on a wall.⁷² Leonard, testifying as part of the prosecution's case, was permitted to opine that the

⁶⁸ Solan & Tiersma, *supra* note 18.

⁶⁹ Gen. Elec. Co. v. Joiner, 522 U.S. 136, 139 (1997).

⁷⁰ People v. Coleman, No. 09-CF-50 (Ill. Cir. Ct. Monroe Cnty. May 5, 2011).

⁷¹ Nicholas Pistor, *Jurors Shown Images of Coleman's Dead Family as Trial Starts*, ST. LOUIS POST-DISPATCH (Apr. 25, 2011, 2:15 PM), http://www.stltoday.com/news/local/Illinois/jurors-shown-images-of-coleman-s-dead-family-as-trial-article_c293b985-a54c-5be1-8540-5c4b7a98432d.html.

⁷² *Id.*

stylistic features of the threat letters had enough in common to suggest that all were written by the same individual and that the letters bore enough similarity to the known writings of Coleman that it was a reasonable—but not proven—hypothesis that Coleman wrote them.⁷³ Butters argued that there was not enough evidence to permit one to draw legitimate inferences.⁷⁴ Coleman was convicted.⁷⁵ As an aside, the circumstantial evidence was strong: Coleman had bought the paint that was used to write the threat on the wall, and the other letters were found on his computer.⁷⁶ Thus, the linguistic debate was offered only to dispel the possibility that someone other than Coleman had used his computer. Nonetheless, the Coleman case demonstrates an instance in which courts permitted forensic stylistic analysis after ruling it admissible in an evidentiary hearing.

Also in the press were stories about an expert declaration by Gerald McMenemy, a linguist who specializes in forensic stylistic analysis. McMenemy testified in a case brought by Paul Ceglia against Mark Zuckerberg, claiming that Zuckerberg did not own the Facebook idea.⁷⁷ McMenemy opined that certain emails allegedly written by Zuckerberg were in fact not written by Zuckerberg. The methods he used have drawn criticism from Professor Butters⁷⁸ and some controversy in the press.⁷⁹ Nonetheless, in an opinion issued but not published as of the

⁷³ Report of Proceedings, May 3, 2011, *Coleman*, No. 09-CF-50 (testimony of Robert A. Leonard).

⁷⁴ Report of Proceedings, Apr. 4, 2011, *Coleman*, No. 09-CF-50 (testimony of Ronald Butters, opposing the admission of Leonard as an expert witness).

⁷⁵ Jim Suhr, *Christopher Coleman, Former Marine, Convicted of Killing Wife, 2 Sons*, HUFFINGTON POST (May 6, 2011, 11:15 AM), http://www.huffingtonpost.com/2011/05/06/christopher-coleman-former_n_858460.html.

⁷⁶ *Id.*

⁷⁷ Decl. of Gerald R. McMenemy in Support of Motion for Expedited Discovery, *Ceglia v. Zuckerberg*, No. 1:2010-cv-00569 (W.D.N.Y. July 9, 2010), available at <http://docs.justia.com/cases/federal/district-courts/new-york/nywdce/1:2010cv00569/79861/50/0.pdf?1307100528>.

⁷⁸ Butters, *supra* note 12, at 354–56.

⁷⁹ See Ben Zimmer, *Decoding Your E-Mail Personality*, N.Y. TIMES, July 24, 2011, at SR12.

time of this writing, a federal Magistrate Judge has credited McMEnamin's position⁸⁰ and recommended that the district court dismiss the case against Zuckerberg and Facebook.⁸¹

McMenamin has testified in authorship cases on many occasions, and in fact has written a book that describes his stylistic approach.⁸² But, unless and until the Magistrate Judge's opinion (or another opinion discussing McMEnamin's contribution) in the Facebook litigation is published, he shows up in the federal and state reporters only indirectly, through a former FBI agent and forensic linguist named James Fitzgerald. In a 2010 federal case decided by the District Court in Utah,⁸³ and a case decided by the federal court in New Jersey in 2000,⁸⁴ Fitzgerald testified in a *Daubert* hearing that he was employing McMEnamin's methods in an authorship attribution case.⁸⁵ In both cases the court permitted him to testify as to similarities and differences between the defendant's known writings and those that were in question in the case but not to offer an opinion as to authorship because the methods did not meet the scientific standard required by the Federal Rules of Evidence.⁸⁶

In contrast, an appellate opinion of the New Jersey state courts affirmed a conviction in a case in which Fitzgerald testified for the prosecution that the defendant had written certain anonymous documents, and Carole Chaski testified that there were significant differences between the anonymous documents and those known to be written by the defendant.⁸⁷ The defendant was actually acquitted of having created the

⁸⁰ Ceglia v. Zuckerberg, No. 1:2010-cv-00569, at 118–19 (W.D.N.Y. July 9, 2010), available at <http://assets.sbnation.com/assets/2383813/ceglia-v-zuckerberg-3-26-13.pdf>.

⁸¹ *Id.* at 151.

⁸² GERALD R. MCMENAMIN, *FORENSIC STYLISTICS* (1993).

⁸³ United States v. Zajac, 748 F. Supp. 2d 1340, 1343, 1351 (D. Utah 2010).

⁸⁴ United States v. Van Wyk, 83 F. Supp. 2d 515, 521 (D.N.J. 2000).

⁸⁵ *Id.*

⁸⁶ *Id.* at 523.

⁸⁷ State v. McGuire, 16 A.3d 411, 430 (N.J. Super. Ct. App. Div. 2011).

anonymous documents.⁸⁸ The matter of fact discussion by the appellate court, however, suggests no problem with trial judges admitting expert testimony of forensic linguists, including opinion as to authorship, whether they conduct their work through stylistic comparison or by virtue of algorithms that they have developed. When one party decides to deal with the other side's expert by hiring his own, there will typically be no rejection of either expert. Judges are not likely to exclude a witness absent an objection from the opposing party. Moreover, unless the case results in a published decision, there will be no publicly salient record of the entire episode. In fact, the court in this case mentioned the forensic linguistic testimony only as an aside, since the defendant had been acquitted on the count for which the testimony was offered.

My goal in this discussion is not to criticize the linguists whose methods were at issue in these cases. On the contrary, much of this essay is devoted to suggesting that stylistic analysis is not provably less reliable than the quantitative methods. My hope is that through communication among those who approach the field from different perspectives, it becomes possible to make such methods crisp enough to withstand scrutiny or at least to integrate their most acute insights into quantitative models.

IV. CURRENT TRENDS IN FORENSIC LINGUISTIC AUTHORSHIP ATTRIBUTION

The field appears to be developing to bring a healthy balance between theory and practice in forensic linguistic identification. The basic problem that the field must address is this: as we learned from Noam Chomsky more than a half century ago, language is a creative cognitive function.⁸⁹ By that, I do not mean to say that we can all be poets if we wish. What "creative" in this context means is that we can produce and understand infinitely many utterances because the rules of a recursive grammar that we have internalized in our minds, mostly as young children, combined with a rich vocabulary, give

⁸⁸ *Id.* at 420.

⁸⁹ NOAM CHOMSKY, LANGUAGE AND MIND 88-91 (3d ed. 2006).

us the capacity to do this. And we indeed use the capacity. As Malcolm Coulthard has pointed out, based on the work of some computational linguists,⁹⁰ if you highlight any ten-word string from any document, and then paste that string into a search engine (e.g., Google) or database (e.g., Lexis) window, you will find that your string is unique among the vast array of documents available.⁹¹ This shows enormous intra-authorial variation in the use of language. We really do not say things the same way twice. What this all means is that the challenge facing authorship attribution research is to discover ways in which the variation in the use of language between authors is demonstrably greater than the variation within a particular author and to focus on the presence or absence of these distinguishing features.

As noted, there seems to be a methodological, and for that matter, cultural division between computational and stylistic analysts, with the former approaching authorship attribution problems in a manner more consistent with conventional views of expert testimony than the latter but with no evidence that either approach does a better job. Three things are likely to bring about a convergence. First, I expect that if the computational approaches reach some level of maturity and consensus, they will be far more often accepted by the courts than those approaches based largely on intuition—even excellent intuition based on experience and learning. Computational linguistics is an established subfield of linguistic inquiry. Thus, it really is possible to apply well-known linguistic techniques to a forensic problem and to ask whether the application in the forensic setting meets the standards of the field itself.⁹² By the

⁹⁰ Fintan Culwin & Mike Child, *Optimising and Automating the Choice of Search Strings when Investigating Possible Plagiarism*, PROC. 4TH INT'L PLAGIARISM CONF., 2010. This article demonstrates that the procedure generally works with as few as six consecutive words. In earlier work, Coulthard had proposed ten words.

⁹¹ Krzysztof Kredens & Malcolm Coulthard, *Corpus Linguistics in Authorship Identification*, in THE OXFORD HANDBOOK OF LANGUAGE AND LAW, *supra* note 17, at 504, 513–14. Of course, the string may also be a direct quotation of a unique expression. No doubt the text of the U.S. Constitution appears on the web in many sites.

⁹² See Argamon & Koppel, *supra* note 17; Carole E. Chaski, *Best*

same token, computational linguists and computer scientists are accustomed to testing their algorithms to see how well they work and reporting the rate of error. One conference, for example, requires the submission of an answer to an authorship attribution task as a criterion for participation.⁹³ These procedures are consistent with contemporary views of acceptable scientific evidence.⁹⁴ If judges, at least in the United States, begin to accept authorship identification as a routine matter precisely because it is transparently algorithmic, with identifiable rates of error, then such work will become the standard.

Second, it will be incumbent upon those whose work is more intuitively stylistic to demonstrate its scientific underpinnings. This can be accomplished by incorporating stylistic features into the computational algorithms being developed by computational linguists and computer scientists. The insightful observations of stylistic analysts that take advantage of such nuances as word choice, punctuation, and spelling errors can be used to expand the range of factors that computer scientists include in their models, with the potential of adding power, even if only incrementally. In fact, this is already occurring.⁹⁵ Progress can also be made through the creation of models that demonstrate the efficacy of stylistic analysis as its own approach. Tim Grant, in his article in this volume, has taken a significant step in that

Practices and Admissibility of Forensic Author Identification, 21 J.L. & POL'Y 333 (2013); Patrick Juola, *Stylometry and Immigration: A Case Study*, 21 J.L. & POL'Y 287 (2013); Koppel et al., *supra* note 17; Stamatatos, *supra* note 17; Carl Vogel, *Attribution of Mutual Understanding*, 21 J.L. & POL'Y 377 (2013); *see also* Chaski, *supra* note 17, at 489 (summarizing the state of authorship attribution); Juola, *supra* note 17 (summarizing the state of authorship attribution); Antonio Rico-Sulayes, *Statistical Authorship Attribution of Mexican Drug Trafficking Online Forum Posts*, 18 J. SPEECH LANGUAGE & L. 53 (2011).

⁹³ The PAN Lab, held in conjunction with the CLEF conference, is referred to in the field as PAN/CLEF. *Plagiarism Detection*, PAN, <http://pan.webis.de/> (last visited Feb. 1, 2013).

⁹⁴ *See* SOLAN & TIERSMA, *supra* note 18, at 29–32.

⁹⁵ *See generally* Argamon & Koppel, *supra* note 17 (discussing the importance of language style in authorship attribution); Juola, *supra* note 92 (discussing statistical linguistics to analyze one's writing style against an ad hoc collection of distractor authors).

direction.⁹⁶ This will no doubt require far more collaboration between the forensic linguistic community and those adept at statistical modeling. Such collaboration has not been adequate to date. Many of the questions that Butters raises (such as how large an effect must be for it to merit evidentiary weight) are answered automatically in computational settings by virtue of the models employed. Similarly, for those engaged in the promising methods of using linguistic corpora as reference sets in authorship attribution,⁹⁷ collaboration with modelers will become a necessity.

In the short run, proficiency testing may substitute for the development of methods, although this remains a second-best solution because it does not eliminate the potential for bias in actual case work. Those linguists who demonstrate their ability to identify authorship correctly in a series of problems presented to them would be credited in court for their proven prowess irrespective of whether their methods are replicable. The PAN/CLEF conference mentioned above,⁹⁸ in which computer scientists test their methods against each other as a prerequisite for participation in the event, might be expanded to include those who analyze authorship cases through stylistic comparison. This is an especially appealing solution in that it may well be the case that the stylistic analysts are highly accurate in their conclusions. Personal communication with some of the organizers of this conference suggests that they would be enthusiastic about such participation. It is also important, of course, that those whose work depends on algorithms be willing to subject their models to testing of the sort described above. Self-proclaimed excellence is no more scientific when asserted by computer scientists than when asserted by stylistic analysts.

Third, and related to my second point, some computer scientists and some linguists have taken to looking at very large sets of features, largely stylistic markers, not concerning

⁹⁶ See Tim Grant, *TXT4N6: Describing and Measuring Consistency and Distinctiveness in the Analysis of SMS Text Messages*, 21 J.L. & POL'Y 467 (2013).

⁹⁷ See Kredens & Coulthard, *supra* note 91, at 504–05.

⁹⁸ See *supra* note 93.

themselves with which features appear relevant in a particular case, as long as enough do. One such research project uses cluster analysis to sort through a large corpus of emails and to at least reduce the pool of potential authors from a great many to a few with considerable accuracy.⁹⁹ Tim Grant's work also makes use of "a basket of features," concerning itself more with the number of features in common than with whether a particular feature or kind of feature is likely to be predictive across cases.¹⁰⁰

Whether I am right or wrong in predicting the subsequent direction of the field, I firmly believe that far more collaboration among scholars with different areas of expertise is absolutely essential. And I would like nothing more than to see a significant increase in collaboration notwithstanding any conflicts experienced by the players and without a precipitating crisis.

⁹⁹ Farkhund Iqbal et al., *Mining Writeprints from Anonymous E-mails for Forensic Investigation*, 7 DIGITAL INVESTIGATION 56, 56 (2010).

¹⁰⁰ Grant, *supra* note 96.