

The Determination of Authorship from a Homogenous Group of Writers

Marie E. Durina¹ and Michael P. Caligiuri, Ph.D.²

¹Forensic Document Examiner, San Diego Sheriff's Crime Laboratory, 5255 Mt. Etna Drive, San Diego, CA 92117

²Professor of Psychiatry and Director, Clinical Research Protections Program, University of California, San Diego School of Medicine, 3350 La Jolla Village Drive, San Diego, CA 92161-116A

Forensic document examiners may encounter challenges when examining specimens from homogeneous writing populations and may need to identify potential sources of errors when rendering conclusions of authorship on such writings. A research project was conducted in which samples of writing were obtained from 52 adult writers who grew up in the same neighborhood, were taught the same copybook style, at the same Catholic elementary school, by the same teachers, approximately 4 decades ago. The specimen writings were subsequently examined and compared by 49 forensic document examiners throughout the world. The examiners rendered conclusions of authorship on the writings and submitted their conclusions for evaluation of accuracy. The results of the study offered evidence to support that there is a high degree of inter-writer variation among writers, even in populations where the driving forces for variation were low; and among these homogeneous writing populations, forensic document examiners were able to extract features from the writing samples that enable them to attribute authorship. The study examined effects of certain factors such as examiner experience, geographic location of examiner, and length of the questioned documents and how these factors affected accuracy. The research addressed criticisms that earlier studies on the individuality of handwriting did not include populations from homogeneous writing communities and relied on computer analysis of handwriting rather than on human examiners.

Overview

Acceptance of testimony from handwriting experts seems to have occurred in the United States as early as 1812 (*Sauve v Dawson*, 2 Martin 202); and the passage of U.S. Statute 683, Chapter 79 in 1913 further expanded the admissibility of such testimony. In 1910, Albert Osborn published 1 of the first textbooks on the subject, *Questioned Documents*. Osborn, who began his career as a teacher of penmanship, is now considered to be the founding father of forensic document examination in the United States.

A most important case in the U.S. that was instrumental in bringing the profession of handwriting examination into the public eye was the 1932 trial of Bruno Hauptmann for the kidnapping and murder of the infant son of Charles Lindbergh. At this trial, 8 forensic document examiners (FDEs) testified and identified Hauptmann as the writer of all of the kidnap/ransom letters. Touted as "the crime of the century" that became "the

trial of the century" (Haring 1937), it is this case that, arguably, made forensic document examination famous. In this case, and in many others for years afterward, FDEs held to Osborn's basic tenets regarding handwriting examination and comparison; i.e., that no 2 persons write exactly the same way, and each adult writer's handwriting contains features that make it unique and identifiable to that particular writer (Osborn 1929). Until the early 1990s, expert testimony of this type given by forensic document examiners was seldom disputed.

Over the past 20 years, however, forensic document examination, particularly the examination of handwriting, has come under scrutiny. In 1989, the publication of "Handwriting Identification Expertise, Exorcism of Ignorance" by 3 law professors in the *Pennsylvania Law Review* (Saks and others 1989) criticized the reliability and validity of forensic document examination. The authors alleged that there had never been any scientific proof provided by experts that shows they can

actually do what they say they can do, nor had there been any empirical studies performed to prove that handwriting is, in fact, unique to only 1 specific individual. Rulings in court cases such as *Daubert v Merrell Dow Pharmaceuticals* (1993), *U.S. v Starzecpyzel* (1995), *General Electric Co. v Joiner* (1996), and *Kumho Tire Co., Ltd. v Carmichael* (1999) ushered in a new era for testimony by FDEs and raised standards of admissibility of evidence. In court, challenges to the scientific basis for handwriting identification began to arise, and the goal of expert critics was to keep juries from hearing testimony about results they considered to be scientifically unreliable—results that would identify a single individual as the source of the handwriting evidence.

In some ways, the critics were correct. For a time there was, as they claimed, a dearth of empirical evidence to support the claims of expertise made by examiners or to support individuality of handwriting. Challenges raised as a result of cases such as *Daubert*, *Starzecpyzel*, *Joiner*, and *Kumho* required that FDEs be able to offer proof that their science was a valid one. To answer these challenges, FDEs began participating in studies assessing their abilities to compare certain features of handwriting and to determine authorship of that handwriting in a valid manner. Several of these studies were conducted by Dr. Moshe Kam of Drexel University and provided supporting evidence that document examiners can examine handwriting with an error rate significantly lower than that of lay persons (Kam and others 1994, 1997, 1998, 2001). For example, in Dr. Kam's large-scale study in 1997, FDEs incorrectly matched 6.5% of the handwriting samples, while laypersons incorrectly matched 38.5%. In his 2001 study involving signatures, the error rate for FDEs who determined genuine signatures to be nongenuine was 7.05%, compared with an error rate of 26.1% for laypersons who performed the same examinations.

An additional study regarding handwriting identification was conducted at State University of New York in Buffalo by Dr. Sargur Srihari, "Individuality of Handwriting" (2002). This research was conducted using a software program known as CEDAR-FOX (rather than human FDEs) to analyze similarities and differences between the handwriting samples taken from over 1,500 individuals in 6 different states throughout the U.S. The CEDAR-FOX program facilitated analysis of the handwriting samples using "computational features" that were determined

algorithmically. These included 11 macro-features of the writing that were categorized into Measures of Pen Pressure, Measures of Writing Movement, Measures of Stroke Formation, Slant, and Word Proportion, as well as 2 micro-features that corresponded to automatic character recognition. Handwriting examinations using the CEDAR-FOX program were quite unlike those performed by human FDEs, who use a multitude of qualitative and quantitative features when examining handwriting. In fact, the textbook, *Handwriting Identification: Facts and Fundamentals* (Huber and Headrick 1999), states that FDEs take into account at least 21 discriminating elements of handwriting. The CEDAR-FOX software did, however, provide an accurate analysis of the handwriting samples. The results of Dr. Srihari's study supported the theory of uniqueness in handwriting and caused him to conclude, "...we were able to validate handwriting individuality with a 95% confidence. By considering finer features, we should be able to make this conclusion with a near 100% confidence."

After Dr. Srihari's study, 2 rebuttals were published, both authored by Michael J. Saks and both critical of the study. The 1st critical publication was "Commentary on: The Individuality of Handwriting" (2003), a discussion only of Srihari's study. The other, an essay entitled, "The Individualization Fallacy in Forensic Science Evidence" (Saks and Koehler 2008), reiterated the criticisms of Srihari's study that Saks had expressed earlier and targeted other forensic disciplines as well. In these publications, when addressing the issues of the Srihari study, Professor Saks contended that inter-writer variation would be expected because Dr. Srihari collected specimen writing from 1,500 individuals throughout the U.S. rather than from smaller "homogeneous writing communities" (i.e., where all of the writers had been trained to write using the same methods).

A 2nd criticism was that the source document used by the writers to prepare the writing sample for Srihari's study contained an abundance of letters and letter combinations and therefore "artificially maximizes the ability to distinguish writers." In regard to the size of the source document, Professor Saks went on to say, "Were the writing sample to decrease in the amount of writing diversity (and begin to approximate more forensically typical writing), the ability to distinguish writers would decline further" (2003).

A 3rd criticism had to do with the fact that “the study involved no human examiners, only computer-based pattern recognition;” therefore, Saks alleged, “...whatever the study found, it could not tell us how well or poorly humans could distinguish one writer from another.”

This research project was undertaken in an attempt to address these 3 criticisms. The present study offered some supporting evidence that, while we can expect to see certain similarities in the writing of persons trained under the same handwriting system, in the same school, and by many of the same teachers, each writer will have identifiable features that will be discernible by FDEs and enable them to attribute authorship with a high degree of confidence. This study tested 2 hypotheses: 1) that experience level and geographic location of the practice of document examiners affect their accuracy when determining authorship from an homogeneous group of writers; and 2) that characteristics of the samples, such as the number of words in the document, affect the examiners' accuracy in determining authorship.

Methodology

Source of Handwriting Samples

From the 1920s until about 1974, the Palmer Method of Business Writing was the dominant handwriting system taught in the United States. This method, initially developed by Austin Palmer in 1884, emphasized proper body, shoulder, and hand movements, combined with stringent teaching methods (Palmer 1935). Good handwriting was accomplished by a series of daily writing drills, with an emphasis on legibility and speed. FDEs often refer to handwriting systems such as the Palmer method as a “copybook style,” where the student is encouraged to practice writing by imitating the letter forms set forth in the copybook and to conform to them. Palmer's methods did not allow for a high degree of nonconformity. In fact, in *Handwriting in America* (1996), Tamara Thornton wrote, “Certainly nothing in Palmerian teaching methods—the calisthenics, the commands, the counting, the drills—encouraged students to assert their individuality through stylistic innovation.”

One of the Catholic elementary schools where the Palmer method was diligently taught and strictly enforced by the nuns of the order of the Sisters of St. Joseph was Our Lady of Perpetual Help Elementary School (OLPH) in Brooklyn, New York. There, 1 of the authors learned penmanship in the early 1960s, along with her 4 siblings and

their childhood friends and classmates. Many of the teachers at that school learned to write with the same methods when they were children years earlier. It was from this pool of persons that the handwriting specimens used in this research study were obtained.

Samples of handwriting were obtained from former students—adults in their 40s through 60s at the time of the study—who learned to write at the same Catholic elementary school during the period between 1950 and 1969, a time during which the Palmer method of penmanship was taught there. In addition, some samples were obtained from the Catholic nuns of the order of the Sisters of St. Joseph, some of them over 80 years old at the time of this research. These nuns had taught many of the same students to write using the Palmer method and had themselves learned to write using the Palmer system as children. The students were trained under the Palmer system in the period between 1950 and 1969, and the nuns' training occurred as early as 1926. All of the writers lived in Brooklyn at the time of their training. At the time of this research, all writers (except 1, who is an author of this paper) resided in the New York Tri-State area, encompassing the populated areas in the states of New Jersey, New York, and Connecticut.

Handwriting specimen forms and a single-page questionnaire were sent via e-mail to addresses obtained from the school's website for former OLPH Elementary School students. The current school principal was contacted via U.S. mail. This resulted in contacts with a former principal of the school, a retired nun of the order of the Sisters of St. Joseph who had taught at the school during the 1960s and 1970s. Several more Sisters of St. Joseph who previously taught at the school were contacted via U.S. mail at St. Joseph's Convent in Brentwood, New York, a convent for retired nuns. The youngest specimen writers were former students who were 48 years old at the time of the study; and the oldest was a retired teacher and nun, 88 years old when this research was undertaken. Median age of the writers was 52 years.

All recipients received the same information—a letter briefly describing the study and a questionnaire asking gender, handedness, and dates of attendance at the school and asking for comments about memories they may have had about learning handwriting as children. The comments section was marked “optional.”

The handwriting specimens consisted of each writer writing 1 sample of “The London Letter,”

a source document containing 85 words, historically used by FDEs to obtain specimen writing (Appendix A). The text of this letter appeared on the printed form, followed by ruled blank spacing. A total of 52 responses were returned from former students and teachers. Of the former students, 34 writers (65%) were females, and 6 writers (12%) were males. All of the former students attended OLPH Elementary School and learned to write beginning in the 3rd grade during the years between 1950 and 1969. Additionally, 12 writing specimens (23%) were received from Catholic nuns of the Order of the Sisters of St. Joseph who had taught at the school and/or attended it years previously. These nuns learned the Palmer system of handwriting during the years between 1927 and 1941. Of all writers, 2 (both females) were left-handed, and 1 was ambidextrous. It is noteworthy that the ambidextrous writer was an elderly nun who remarked in her comments that she had not been permitted to write with her left hand as a child.

Document Examiners – Analyses and Comparison of Samples

The samples of The London Letter, 1 from each writer, became the known specimens (Ks) and were marked as K1 through K52. The optional comments were made by only 43 writers on pages separate from The London Letter specimens. These comment pages were separated from the known specimens, became the questioned documents (Qs), and were randomly marked Q1 through Q43. Over 50 qualified FDEs, practicing in both government forensic laboratories and private practice, were notified about the study by e-mail and agreed in their replies to examine and compare the questioned and known writing samples in an attempt to determine authorship. The writing samples were then photocopied on a black-and-white Canon photocopy machine, and the original documents were also scanned into PDF files at a resolution of 600 dpi. These photocopies and PDF images were then provided by mail to examiners in the United States, Canada, Australia, the Netherlands, and Finland, who agreed to examine and compare the documents and to record their answers on the answer sheet provided.

It is important to note that the answer sheets did not provide for the rendering of opinions that were inconclusive or qualified. The answer sheets allowed only for definitive responses; these are sometimes referred to by document examiners as

“forced call” opinions. In this regard, the study did not resemble real casework where there can be up to 9 levels of opinions (ASTM International Standard E1658-04). Many examiners made notes on the answer sheets indicating that if certain samples had been encountered in real casework, their opinions would have been far more conservative than those they were expressing for the purposes of this study. In these cases, examiners noted they would have rendered conclusions that were qualified rather than definitive; or they would have reported that their findings were inconclusive, meaning they could neither include nor exclude a particular writer as the author of the questioned document. The examiners who noted that their opinions would have been qualified or inconclusive, had they been permitted to render such opinions, made these notations for many of the same specimens. Completed answer sheets were received from 49 FDEs.

Data Analysis

A 2-way analysis of variance (ANOVA) was used to evaluate main effects of years of experience and country of origin and of the 49 examiners on response accuracy. The experience factor consisted of 2 levels (equal to or below 10 years and greater than 10 years of experience). The country factor also consisted of 2 levels of examiner geographic location of practice (North America and non-North America). In addition, we employed multiple regression analyses to examine whether years of experience and length of the questioned documents (in number of words) accounted for significant variability in examiner response.

Results

Overall Accuracy

The answer sheets collected from the 49 participating FDEs were scored to determine the number of correct and incorrect responses. The respondents were asked to complete a short questionnaire with their answer sheets documenting their years of experience and state whether they worked alone or with other examiners in completing the answer sheets. Several examiners provided additional comments about their participation in the study. Of the 49 participants, 15 reported they engaged in a peer review process where their responses were corroborated by at least 1 other examiner; and 1 examiner reported that a limited review was conducted of the examiner's answer sheet by a 2nd examiner. There were 16 examiners

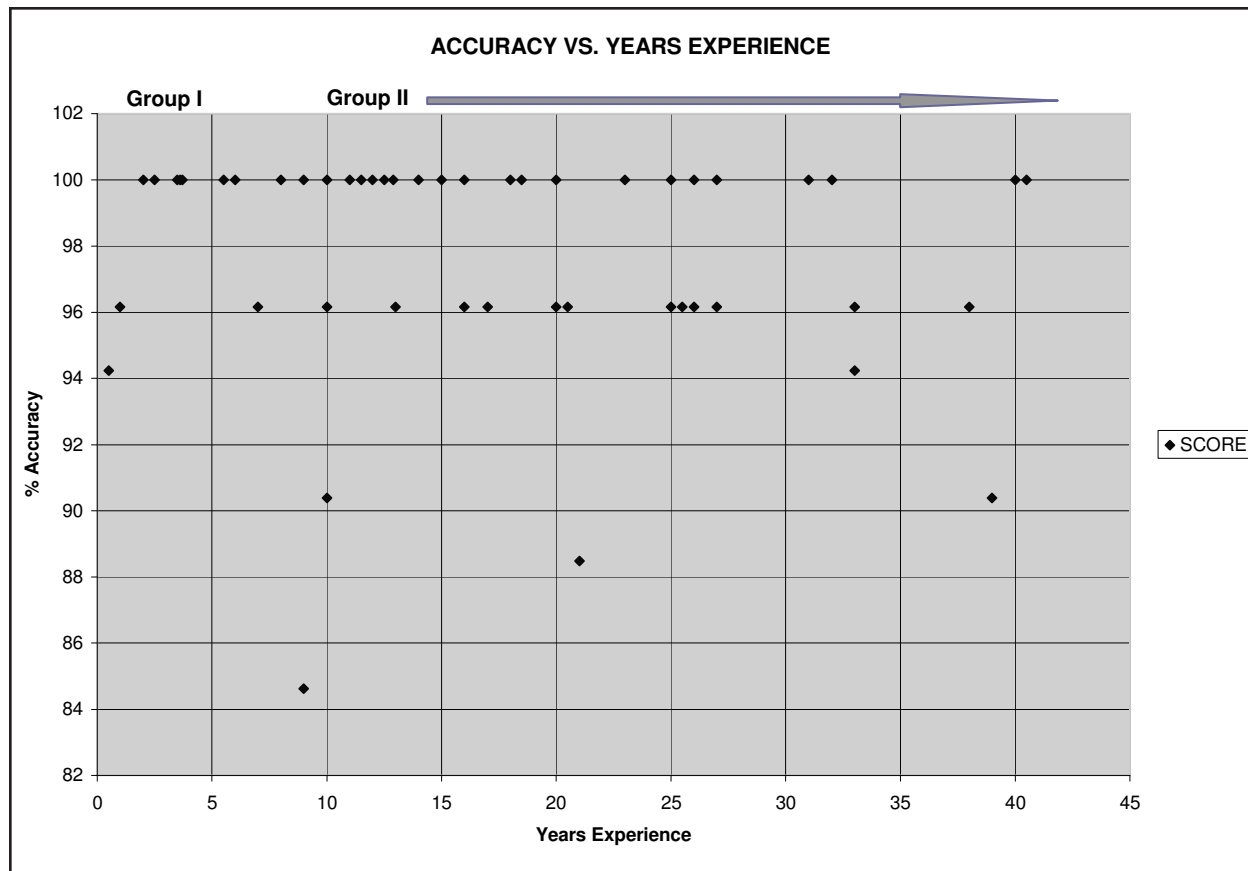


Figure 1. Results of examiners, shown with corresponding number of years of experience. Average group score was 98%.

with 10 years or less of experience (Group I) and 33 examiners with greater than 10 years of experience (Group II).

With a perfect score valued at 100 points (100% correct pairings of Qs and Ks), and a total of 52 selections possible from the pool of known writing specimens, this meant each of the 52 possible answers had a value of 1.923 points. For each incorrect response, or for answers left blank, 1.923 points were deducted. For example, if 1 of the Q documents was matched with a K writer who was not the author of that document, this resulted in a deduction of 1.923 points and resulted in a score of 98.08%. If a K writer was selected as an author for 1 of the Qs, but this K writer was 1 of the 9 writers who did not author an associated Q document, this resulted in an additional deduction of 1.923 points (i.e., 2 incorrect responses). For reporting purposes, this score was rounded up or down to the nearest whole number. For example,

2 incorrect choices resulted in a score of 96.154%, which was recorded as a score of 96%.

Even with the confines that were imposed upon the examiners in making “forced call” conclusions, the overall results were quite favorable. As a group, the average score for all participating examiners was 98%. Twenty-nine examiners scored 100%; 14 examiners scored 96%; 2 examiners scored 94%; 2 examiners scored 90%; 1 examiner scored 88%; and 1 examiner scored 85% (Figure 1).

No meaningful relationship was found between years of experience and correct associations made between questioned and known specimens. For example, some very experienced examiners achieved perfect scores: 19 examiners in Group II. Yet, the other 14 in this group did not. Some less experienced examiners, 10 in Group I, also achieved perfect scores, while 6 did not. These findings are in line with results of a handwriting

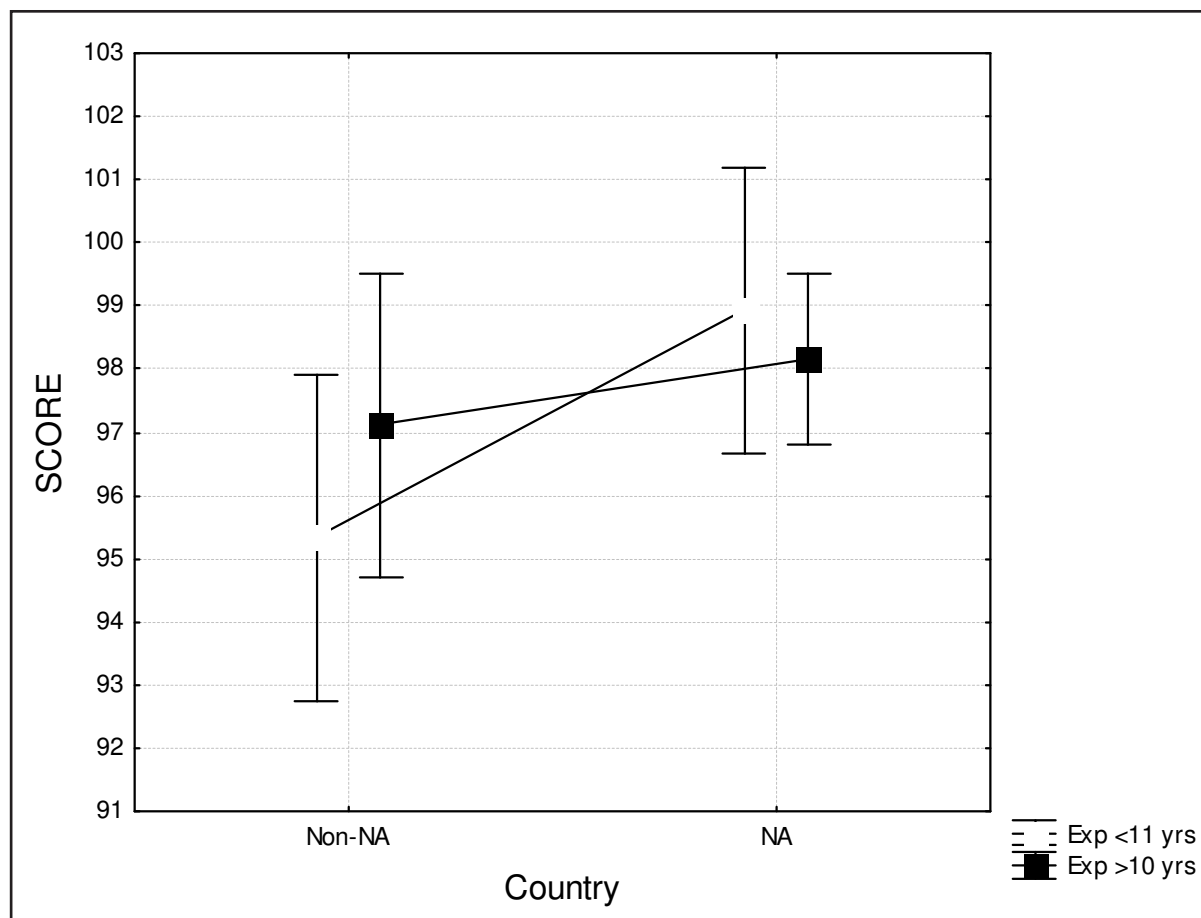


Figure 2. Mean score (with 95th confidence intervals) for 2 groups of examiners (separated by years of experience) from North America (NA) or non-NA countries.

study reported in 2002, where “No relationship was found for experience and number of correct opinions made” (Sita and others). It is noteworthy that the 15 examiners who reported engaging in a full “peer review” process achieved perfect scores every time. Although peer review was not necessary to achieve a perfect score, and, in fact, 14 of the 29 examiners who scored 100% did *not* report that they engaged in peer review, the peer review process appeared to offer some benefit. Results support the use of peer review as a valuable tool in real casework, when it is available.

Associations between Examiner Classifications and Accuracy

Figure 2 shows the results of the 2-way ANOVA. These results indicated that examiners from outside North America (non-NA) had higher error rates than North American examiners (NA) when

controlling for experience level ($F_{1,45}=4.50$; $p<0.05$); whereas, years of experience was not significant ($F_{1,45}=0.21$; $p>0.10$). The interaction between experience and country code was not significant ($F_{1,45}=1.38$; $p>0.10$), suggesting that the difference in response accuracy between examiners from NA versus those from non-NA was independent of years of experience.

Figure 3 shows the relationship between document length and response accuracy. The length of each of the questioned documents (Q1 through Q43) was determined by counting the number of words in an effort to determine if there was a significant correlation between the length of the questioned document (document length in words) and examiner accuracy (score). There was a significant correlation between document length and accuracy scores. We found a nonlinear fit between accuracy and document length of the questioned

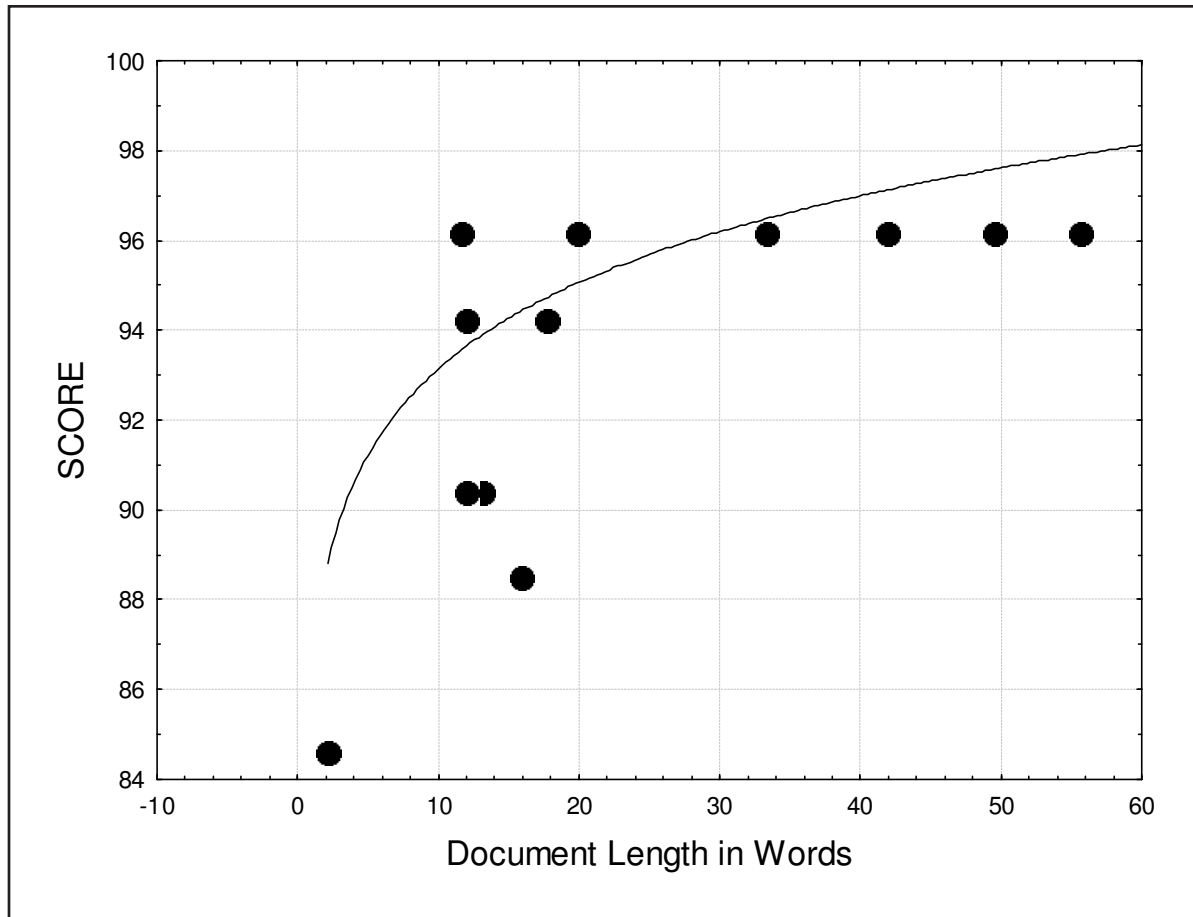


Figure 3. Nonlinear relationship between accuracy score and document length (in number of words) for 20 examiners scoring less than 98% on the handwriting matching task.

documents. Correlational analyses indicated a significant relationship between the log of the document length and response accuracy ($r = 0.60$; $p < 0.01$), with the shorter length documents being associated with lower accuracy scores. Based on this analysis, documents with document lengths less than 20 words were associated with greater examiner error scores than documents having word lengths of greater than 20 words.

Multiple Regression Models

Stepwise multiple regression procedures were used to account for the variability in examiner response using 3 independent variables: the log of document length, years experience, and county code. Table 1 shows the results of the 2- and 3-factor regression models. Each model included data from the 20 examiners who scored less than 100% accuracy. The result for R^2 indicates the

amount of variability in response error that could be accounted for by the factors of the model.

As can be seen from Table 1, models that included document length account for a statistically significant portion of the variability in response error. Adding years of experience to a 2-factor model consisting of country code and document length increased the strength of the model by 9% and accounted for over 50% of the variability in response error.

Problematic Specimens

As can be seen in Table 2, a total of 13 questioned documents were found to be problematic for examiners. Incorrect pairings occurred, but these were not always paired with the same incorrect partner. Some incorrect pairings occurred when the Q document contained a very limited quantity of handwriting.

	Factor 1	Factor 2	Factor 3	Statistic (p-value)
Model 1	Country Code	Years Experience	NA	R2 = 0.13 (p>0.10)
Model 2	Years Experience	Log Document Length	NA	R2 = 0.38 (p<0.05)
Model 3	Country Code	Log Document Length	NA	R2 = 0.44 (p<0.01)
Model 4	Country Code	Log Document Length	Years Experience	R2 = 0.53 (p<0.001)

Table 1. Results from multiple regression models predicting response accuracy.

Questioned documents Q32 and Q28 caused the highest number of incorrect pairings. They are shown with their associated known specimens (Figures 4 and 5).

As can be seen in Table 3, a total of 11 known specimens were found to be problematic for examiners. Every K document contained 85 words. These were not always paired with the same incorrect partner.

Known specimen documents K32, K9, and K48 caused the highest number of incorrect pairings (Figures 6 and 7).

Many examiners commented on the structure of the study and observations they made while examining the documents. As previously stated, many examiners expressed their uneasiness in rendering definitive conclusions about questioned documents that contained limited writing. Another common thread was that the examiners felt they were able to determine authorship without a great deal of difficulty for the majority of the specimens but found analysis of the small group of remaining specimens to be very challenging and time-consuming. Several examiners noted

Questioned Document Number	Number of Incorrect Pairings	Percentage
Q7	1	3%
Q10	1	3%
Q16	1	3%
Q18	1	3%
Q20	1	3%
Q25	5	12%
Q28	8	19%
Q31	2	4%
Q32	11	27%
Q35	1	3%
Q37	1	3%
Q39	6	14%
Q43	1	3%
TOTAL	13	100%

Table 2. Problematic questioned documents.

I would always want to rewrite my notebooks, so I could have the same beautiful handwriting of my teacher (the Nuns is what I really mean). I wish the children

"Our London business is good, but Vienna and Berlin are quiet. Mr. D. Lloyd has gone to Switzerland and I hope for good news. He will be there for a week at 1496 Zermott St. and then

Figures 4a and 4b. Q32 (above) was incorrectly paired on 11 occasions. Its correct association was K40 (below).

they observed different writing styles for some of the writers, where the known specimens of The London Letter appeared to have been more formally and carefully written than the same writer's comments observed on the questioned documents.

Conclusions

The impetus for this study was a desire to address some of the issues noted by Professor Saks in his critique of Dr. Srihari's study on the "Individuality of Handwriting" because that study only involved samples taken from a large, non-homogenous group of writers throughout the U.S. and because it relied on the analysis of handwriting by computers rather than by human examiners.

Forensic document examiners are often faced with the task of determining authorship of a questioned document through comparison to a larger group of known writing specimens. It is hoped that this study offers supporting empirical evidence that there is some degree of inter-writer variation even in homogenous populations where the driving forces for variation are low; and even when inter-writer variation is low and the writing

samples are somewhat limited, FDEs can extract features from handwriting samples which enable them to attribute authorship with a high degree of confidence.

Dr. Srihari's 2002 study in which a computer program was used to analyze writing from a non-homogeneous group of writers concluded, "Based on a few macro-features...from a hand-written document and micro-features...from a few characters, we were able to establish with a 98% confidence that the writer can be identified." Interestingly, in our study, the average score of accuracy for the entire group of 49 participating examiners was also 98%.

Results of our study support that FDEs were able to select fine and subtle features within similar bodies of writing that enabled them to successfully exclude or include writers as potential sources of questioned writing, even among homogenous writing populations. There is also some evidence to support that the process of peer review aided the examiners in these efforts. Multiple regression and ANOVA analyses of the data found no significant correlation between an examiner's years of experience and accuracy. There was, however,

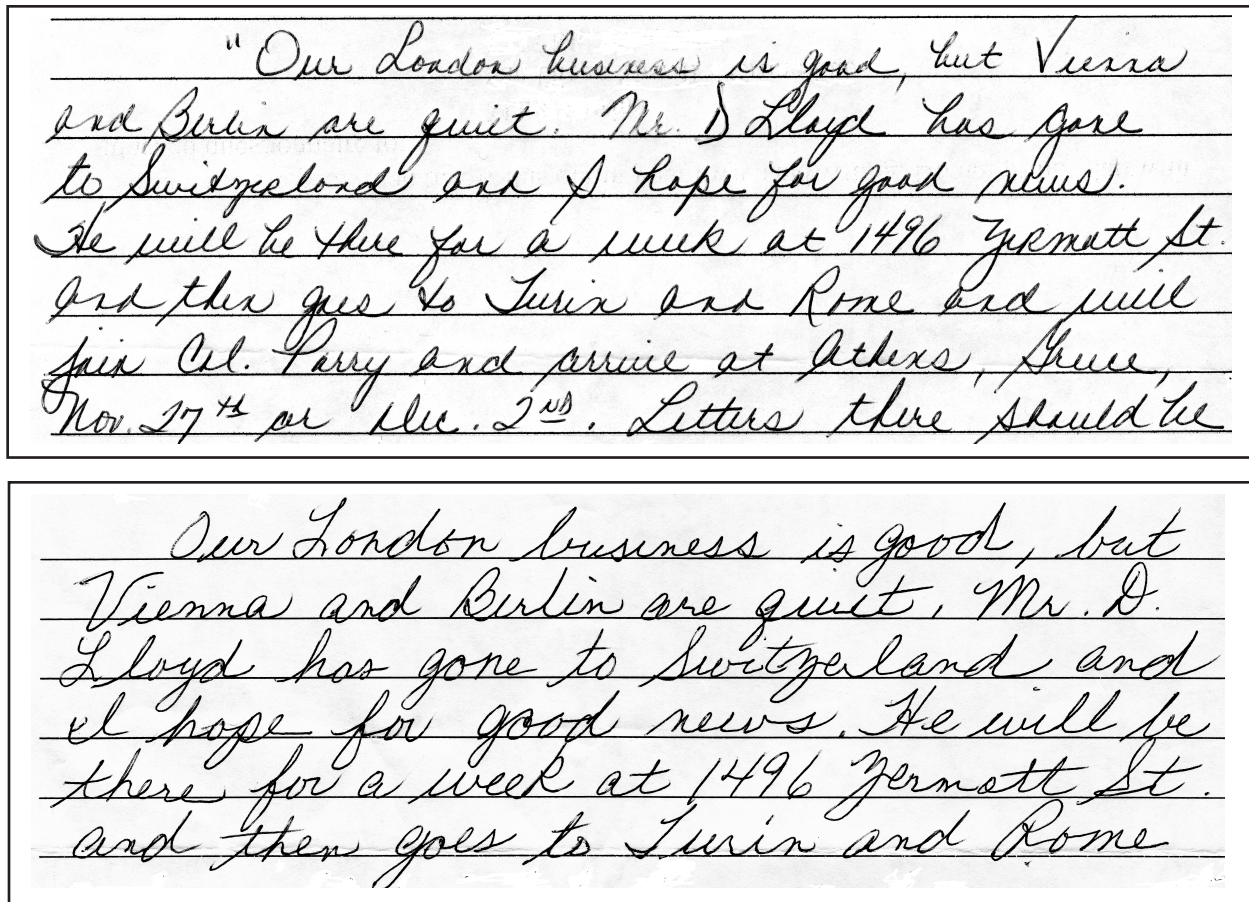
I remember learning to write in script was an arduous task. No matter how hard I tried it was difficult to form the letters according to the teachers instructions. To this day I do not feel confident in my handwriting skills.

"Our London business is good, but Vienna and Berlin are quiet. Mr. D. Lloyd has gone to Switzerland and I hope for good news. He will be there for a week at 1496 Gerrard St. and then goes to Turin and Rome and will join Col. Parry and arrive at Athens, Greece, Nov. 27th or Dec. 2nd.

Figures 5a and 5b. Q28 (above) was incorrectly paired by examiners on 8 occasions. Its correct association was K36 (below).

Questioned Document Number	Number of Incorrect Pairings	Percentage
K9	3	11%
K16	1	4%
K18	2	7%
K24	1	4%
K28	2	7%
K32	11	40%
K38	1	4%
K39	1	4%
K40	1	4%
K48	3	11%
K50	1	4%
TOTAL	11	100%

Table 3. Problematic known specimens.



Figures 6a and 6b. K32 (above) and K9 (below) were not associated with any of the Q documents; however, K32 was incorrectly paired 11 times, and K9 was incorrectly paired 3 times.

significant correlation between accuracy score and examiner country of practice, with examiners located in North America having higher accuracy scores. There was also significant correlation between the length of the questioned document (number of words) and accuracy scores, with the documents of shorter length (approximately 20 words or less) being associated with lower accuracy scores. These findings offer some evidence to support that forensic examiners should continue to exercise caution when examining foreign writings, as well as when examining writings of limited word length.

This study was by no means all-encompassing, as there were limitations. Participation in the study was limited to a total of 49 FDEs. Participating FDEs worked with photocopies and PDF (digital) images rather than original documents. The writers from whom the specimens were obtained were not observed when writing,

and no data was obtained from them regarding any medical conditions, physical limitations, and mental states. All samples, except 1, were written with ballpoint pens, but brands used are unknown.

This study did not allow for qualified opinions to be rendered and, in this regard, did not mimic real casework. Yet in other ways the study did closely replicate conditions faced by FDEs in the performance of their duties. For example, in this study 13 of the questioned documents consisted of writing specimens limited to 30 written words or less. All of the questioned documents varied in length and were not directly comparable to the known specimens. The known specimens were limited to 1 single specimen of the same paragraph of writing (85 words) from each writer and, therefore, may not have represented each writer's full range of natural variation. Finally, many of the specimens displayed a difference in writing

"Our London business is good, but Vienna and Berlin are quiet. Mr. D. Lloyd has gone to Switzerland and I hope for good news. He will be there for a week at 1496 Zermott St. and then goes to Turin and Rome and will join Col. Parry and arrive at Athens, Greece, Nov. 27th or Dec. 2nd.

both. I was left handed. The only right hand was changed to accomodate the teacher. Changing to writing with my right hand was very difficult.

Figures 7a and 7b. K48 (above) was incorrectly paired 3 times. Its correct association was Q39 (below).

style between each writer's specimen K and the associated Q, where the Q was often written in a casual manner, while the associated K specimen of The London Letter was written in that writer's more formal copybook style. The need to examine limited comparable writing, writing executed by the same writer for formal and informal purposes, or writing executed in styles that may not encompass each writer's full range of variation are challenges document examiners encounter frequently in their normal casework.

Future studies may include the use of computer analysis of the specimens and comparing those results with the results recorded by FDEs, using computer software to examine frequency of

occurrence of certain writing features within the samples and comparing results recorded by FDEs with those of lay persons who may be willing to participate in such a project.

Acknowledgments

The authors wish to thank the former students of Our Lady of Perpetual Help Elementary School in Brooklyn, and the Sisters of St. Joseph who taught there, for the donation of their handwriting specimens for this study. Special thanks to the retired Sisters from the St. Joseph Convent in Brentwood, Long Island; to retired OLPH Elementary School principal Sister Madeline

Therese; and to former students Alice Hansen, Bernadette Hoban, and Sheila Pisciotta for their tenacity in soliciting writing samples from other former students and teachers.

The authors are extremely grateful for the participation of the 49 forensic document examiners from Australia, Canada, Finland, the Netherlands, and the United States who undertook the time-consuming examination and comparison of the handwriting samples and who diligently reported their conclusions.

References

- Civil Code of Louisiana ruling, *Sauve v. Dawson*, 2 Martin 202 (1212), Civil Code of Louisiana, as reported in an article entitled Some Historical Aspects of Document Examination, *FBI Law Enforcement Bulletin*, April, 1965.
- U.S. Statute 683, Chapter 79, enacted by Congress on February 26, 1913.
- U.S. Supreme Court ruling, *Daubert v Merrell Dow Pharmaceuticals, Inc.* 509 US 579, 1993.
- U.S. District Court ruling, *United States v Starzecpyzel*, 880 F. Supp. 1027 (S.D.N.Y.), 1995.
- U.S. Supreme Court ruling, *General Electric Co. v Joiner*, 78 F. 3d, 524, 1997.
- U. S. Supreme Court ruling, *Kumho Tire Co. v Patrick Carmichael*, 131 F 3d 1433, 1999.
- ASTM Standard 1658-04, (2004). *Standard Terminology for Expressing Conclusions of Forensic Document Examiners*.
- Haring, J. Vreeland (1937). *The Hand of Hauptmann*, Plainfield, NJ, The Hamer Publishing Company.
- Huber, R.A. and Headrick, A.M. (1999) *Handwriting Identification: Facts and Fundamentals*. Boca Raton, FL, CRC Press.
- Kam, M.; Fielding, G.; and Conn, R. (1997). Writer Identification by Professional Document Examiners, *Journal of Forensic Sciences*, Vol. 42, pp. 778-785.
- Kam, M.; Fielding, G.; and Conn, R. (1998). Effects of Monetary Incentives on Performance of Non-professionals in Document-Examination Proficiency Tests, *Journal of Forensic Sciences*, Vol. 43, pp. 1000-1005.
- Kam, M.; Gummadidala, K.; Fielding, G.; and Conn, R. (2001). Signature Authentication by Forensic Document Examiners, *Journal of Forensic Sciences*, Vol. 43, pp. 884-888.
- Kam, M.; Wetstein, J.; and Conn, R. (1994). Proficiency of Professional Document Examiners in Writer Identification, *Journal of Forensic Sciences*, Vol. 39, pp. 5-14.
- Osborn, A.S. (1929). *Questioned Documents*, Revised Edition. Albany, NY, Boyd Printing Co.
- Palmer, A.N. (1935). *The Palmer Method of Business Writing*, 1935 Edition, New York: The A.N. Palmer Company.
- Saks, M.; Denbeaux, M.; and Risinger, M. (1989). Handwriting Identification Expertise, Exorcism of Ignorance. *Pennsylvania Law Review*, p. 76.
- Saks, M. (2003). Commentary on: The Individuality of Handwriting, Srihari, et al. *Journal of Forensic Sciences*, Vol. 48, pp. 916-17.
- Saks, M. and Kohler, J. (2008). The Individualization Fallacy in Forensic Science Evidence, *Vanderbilt Law Review*, Vol. 61, No. 1, pp. 199-217.
- Sita, J.; Found, B.; and Rogers, D.K. (2002). Forensic Handwriting Examiners' Expertise for Signature Comparison. *Journal of Forensic Science*, Vol. 47, No. 5, pp. 1117-1124.
- Srihari, S.N.; Cha, S.; Arora, H.; and Lee, S. (2002). Individuality of Handwriting, *Journal of Forensic Sciences*, Vol. 47, pp. 856-872.
- Thornton, T.P. (1996). *Handwriting in America, A Cultural History*, New Haven, CN: Yale University Press. pp. 66-70 and 143-166.
- Zimmerman, B.J. and Schunk, Dale H. (2003). *Educational Psychology: a Century of Contributions*, Mahwah, NJ: Lawrence Erlbaum Associates, pp. 10-20.
- <http://www.zanerian.com/Palmer.html>.

Appendix A

Please write the paragraph below this out in your own handwriting. Do not print. Mail with attached questionnaire to:

Marie Durina
San Diego Sheriff's Crime Lab
5255 Mt. Etna Drive
San Diego, CA 92117

HANDWRITING SAMPLE:

"Our London business is good, but Vienna and Berlin are quiet. Mr. D. Lloyd has gone to Switzerland and I hope for good news. He will be there for a week at 1496 Zermott St. and then goes to Turin and Rome and will join Col. Parry and arrive at Athens, Greece, Nov. 27th or Dec. 2nd. Letters there should be addressed: King James Blvd. 3580. We expect Charles E. Fuller Tuesday. Dr. L. McQuaid and Robert Unger, Esq., left on the "Y.X." Express tonight."
