Journal of the American Society of Questioned Document Examiners, Inc.
Volume 19, Number 2, December 2016

Laypersons' Performance in the Determination of Authorship from a Homogenous Group Of Writers*

Marie E. Durina[1] and Michael P. Caligiuri, Ph.D.[2]

[1]Forensic Document Examiner, Riley Welch LaPorte & Associates Forensic Laboratories, PO Box 70, Frankenmuth, MI, 48734-0070, USA
[2]Emeritus Professor, Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive MS 06203, LaJolla, CA, 92093, USA

In 2009, the authors reported a study involving 49 FDEs from 5 countries who attempted to determine authorship of writing samples from 52 writers who had grown up in the same neighborhood, attended the same school, and who had all learned to write using the Palmer method. In that initial study, FDEs were able to successfully attribute authorship with average accuracy scores of approximately 98%. A subsequent study was conducted involving a group of 46 Laypersons who were assigned the same comparison tasks as the FDEs involving the same writing specimens. This secondary study compared accuracy and error rates attained by the Laypersons with those attained by the FDEs. Findings demonstrated that the Laypersons in this study were able to determine authorship with average accuracy scores of approximately 76%. A comparison of error rates between the two groups showed an error rate of approximately 39% for the Laypersons Group, compared to approximately 3% for the FDE Group. Additional findings showed there were profound differences between both groups in the number of writing specimens that were problematic for the examiners, and in the strategies they used in examination and comparison of the handwriting samples.

Introduction

Throughout the 1990s, rulings in the federal cases *Daubert v. Merrell Dow Pharmaceuticals, Inc*. [1], *United States v. Starzecpyzel* [2], *General Electric Co. v. Joiner* [3], and *Kumho Tire Co. v. Patrick Carmichael* [4], resulted in unintended consequences for forensic scientists engaged in the examination and comparison of pattern and impression evidence such as firearms, toolmarks, fingerprints, footwear and tire impressions, and handwriting. Critics of these forensic science disciplines began to surface with ever-increasing frequency, demanding that conclusions ranging from identification to elimination, and the skill of examiners at drawing these conclusions, be supported with evidence from empirical research. It should be required, they professed, that these examiners demonstrate how accurately and reliably they could, in fact, do what they say they could do. With regard to the examination and comparison of handwriting evidence, it was further alleged that there was no research to support that professional examiners possessed such expertise, and that they were no better at this task than laypersons. Some of the most scathing criticism surfaced in a paper published in the *Pennsylvania Law Review* in 1989 [5.] In that article, the authors stated: "If a jury can compare handwriting no worse than proffered "experts", then the expertise does not exist. For any given task, the level of performance of professional document examiners may be no better than that of laypersons." Similar allegations resurfaced in yet another article in the *Iowa Law Review* where this issue was revisited and the authors once again pointed out the lack of research to support the validity of handwriting examination and comparison: "At this juncture there is insufficient information on the contours of both lay and document examiner accuracy to justify rethinking the role of or justification for allowing such expertise" [6.]

*The authors takes full responsibility for this report, for and the design and completion of the research, the writing of the results, and the authenticity and validity of the article.

Several studies have been undertaken since 1994 in an attempt to establish forensic document examiner expertise when compared to that of laypersons; notably by Kam and others [7, 8, 9, 10], by Sita, Found and Rogers [11, 12], and by Bird, Found and Rogers [13.] However, the publication of a report in 2009 entitled "Strengthening Forensic Science in the United States: A Path Forward" by the Committee on Identifying the Needs of the Forensic Sciences Community (often referred to as the "NAS Report") again thrust the comparative forensic disciplines into the spotlight [14.] Although the Committee cited some of the aforementioned research performed in the area of handwriting comparison, and credited these studies as an aid in providing greater understanding of the profession, the issue of professional document examiner expertise was treated with what can be likened to damning with faint praise.

Yet again, the need for additional empirical research was cited. With regard to handwriting examination and comparison, the NAS Report had this to say: "The scientific basis for handwriting comparisons needs to be strengthened. Recent studies have increased our understanding of the individuality and consistency of handwriting and computer studies and suggest that there may be a scientific basis for handwriting comparison, at least in the absence of intentional obfuscation or forgery. Although there has been only limited research to quantify the reliability and replicability of the practices used by professional document examiners, the committee agrees that there may be some value in handwriting analysis" [15.]

Research continued, as did the development of professional standards in forensic document examination by the Scientific Working Group in Document Examination [16] and ASTM [17.] Recently published additional research included an extensive eye-tracking study involving laypersons and professional document examiners conducted by Dr. Merlino of Kentucky State University and funded by NIJ [18], and another study conducted by Kam [19], which involved the examination of simulated and natural handwriting by both laypersons and forensic document examiners. To date, **all** of the aforementioned published studies have supported that professional forensic document examiners possess abilities superior to that of laypersons when examining and comparing handwriting, and have demonstrated the ability of professional forensic document examiners to perform such tasks with a higher degree of accuracy than laypersons. However, handwriting exemplars used in prior studies of lay examiner proficiency were from a diverse group of heterogeneous writers, a design feature which likely contributed to higher accuracy scores among laypersons.

A supplemental study conducted by the author, and now being reported, will examine results obtained by laypersons and compare them to those obtained by professional forensic document examiners who examined the same group of homogeneous writing specimens.

In 2008 and 2009, the authors conducted an initial research project involving handwriting specimens obtained from a homogeneous writing community that was motivated, in part, by prior research by Dr. Srihari on the individuality of handwriting [20.] In the initial research project on homogeneous handwriting reported in 2009, one of the authors (Durina) obtained handwriting specimens from a homogenous group of 52 writers who had attended, or taught at, the same Catholic elementary school in Brooklyn, NY from the period of 1927 through 1969, all of whom had been trained at that school to write using the Palmer method of handwriting [21.] Fifty-two handwritten specimens of an 86-word document known by forensic document examiners as "the London Letter" were obtained from the writers, and these were numbered and labeled as known writing specimens K1 through K52. Additionally, 43 separate writings of varying length (5-225 words) were obtained from all but 9 of the 52 writers in which the writers commented upon their memories of learning to write. Some of these handwritten commentaries were limited in length, meaning they contained fewer than 20 words. The 43 documents with commentaries were separated by the author from their associated K partner, numbered randomly, and labeled as questioned documents Q1 through Q43. All of the documents (43 Qs and 52 Ks) were scanned at 600 dpi and distributed on compact discs, along with photocopies of the documents, to 49 professional forensic document examiners (FDEs) throughout the world. The participating FDEs were tasked with the examination and comparison of the documents in the attempt to determine authorship by "matching" the questioned document with the correct known writing specimen and recording their answers on a corresponding answer sheet. The FDEs were able to perform this task quite successfully, with an overall average accuracy score of 98% for the group. When errors did occur, they occurred

within the confines of a limited group of documents, determined to be the same problematic group of 13 Q documents and 11 K documents. Additional findings showed that the questioned documents containing fewer than 20 words contributed to errors, as did the geographic location of some participating examiners who were not as familiar with the Palmer method of handwriting (typically found in North America.) The results of that study were reported in 2009 in a paper entitled "The Determination of Authorship from a Homogeneous Group of Writers" [22.]

At the time the initial study was conducted, it involved 49 professional FDEs with years of experience ranging from 2 to over 40 years. As only professional forensic document examiners were permitted to participate in the initial study, no assessment could be made at that time with regard to the comparison of the forensic document examiners' performance to that of laypersons who had no prior training or experience in forensic document examination. This paper will discuss the authors' supplemental study conducted in late 2009 in which 46 laypersons were tasked with the examination and comparison of the same dataset of questioned and known documents that were previously examined by the group of 49 professional forensic document examiners.

Results of previous research conducted in this area by others [7-13, 18, 19] support the hypothesis that professional FDEs possess greater expertise than laypersons in the examination and comparison of handwriting. This supplemental study sought to once again test that hypothesis.

Methods and Materials

In October 2009, one of the authors (Durina) was granted access to students enrolled in an Administration of Justice course entitled "Advanced Forensic Photography" at Grossmont College in El Cajon, CA, in the County of San Diego. Prerequisites for this course were the following two courses: Introduction to Administration of Justice, and Forensic Photography (a basic photography course.) All students participating in this research project had previously completed these two courses. In doing so, they had received instruction and training in ethics in forensic science, function and operation of the criminal justice system, courtroom presentation and testimony, and the importance of the role of scientific evidence in helping to determine guilt or innocence in courts of law. In October 2009, the author gave a brief presentation to the students in the Advanced Forensic Photography class, soliciting their possible participation in a research project in which they would be asked to examine and compare cursively handwritten specimens from 52 writers.

The students were not given any training in handwriting examination and comparison, and were not advised of the purpose of the study (i.e. that the results they ultimately submitted would be compared to results submitted previously by professional FDEs who examined the same specimens.) The students were advised that the requirements for their participation were that they would have to be able to read cursive writing, and they could not have received any prior training or experience in forensic document examination. They were further advised to treat this assignment as if it involved actual forensic evidence. They were instructed to take as much time as possible when comparing the handwriting samples and to keep in mind that any errors on their part could, in real life, result in the wrongful conviction of an innocent person or allow a guilty person to go undetected.

As an incentive to solicit participation in the project, students were further advised by their professor that they would be awarded 20 points of extra credit at the end of the term for their participation; **however**, they were also advised that **one point would be deducted for every incorrect response** they recorded on their answer sheets. As there were 52 separate known handwriting specimens involved, they were reminded this could result in a loss of over 20 points should they match more than 20 specimens incorrectly. This incentive is similar to the incentive methods used by Kam et al in their research on layperson and FDE accuracy rates [9, 10.]

A total of 46 students agreed to participate in the project and subsequently submitted completed answer sheets. These 46 students (designated as the Laypersons Group) consisted of 37 females and 9 males who ranged in ages between 19 and 55 years, with a mean (sd) age of 27.61(7.50) years. Each participating student received a packet of photocopies of 52 known documents labeled K1 through K52, and 43 questioned documents labeled Q1 through Q43, a compact disc with digital images of these documents scanned at 600 dpi, an answer sheet for recording their answers, comments, time spent on the assignment, and the following written instructions:

"There are 52 Known Specimens, and 43 Questioned Documents;

All writing is natural. There is no disguised writing.

Each writer wrote only (1) Known Specimen.

Every writer did not write a Questioned document. (There are 9 more Ks than Qs.)

Every Questioned document will have a Known Specimen associated with it.

Extra credit of 20 points awarded for participation, but 1 point will be deducted for each incorrect answer.

You will be given one month to conduct the examinations and complete the answer sheet.

You are allowed to collaborate with fellow students, but NOT with forensic teachers or professionals."

All documents and Answer Sheets provided to the Laypersons Group were the same as those provided to the FDEs in the initial study [23.] Written instructions were also similar to those given to the participating FDE Group, with the exceptions of the statement regarding the extra credit incentive, and the statement allowing collaboration with fellow students, but not with forensic professionals. The students who comprised the Laypersons Group were given over one month to complete their examinations and comparisons and submit their completed answer sheets. The writing specimens, answer sheets, and instructions were distributed to the Laypersons Group on October 2, 2009 with a deadline for submission of the completed answer sheets of November 11, 2009. (This deadline allowed students one complete month to work on this project, while allotting them an additional week to study for their midterm college examinations.)

Results

Overall Accuracy

The 46 completed answer sheets with comments were collected by the author in mid-November 2009, and scored to determine the number of correct and incorrect responses. This information was tabulated and recorded by the authors. Additional information was also recorded, such as age, gender, whether the respondent worked with another person when comparing the writings (i.e. a loosely defined "peer review"), and how long each student took to complete the assignment. Of the 46 participants, only 4 stated they engaged in some form of peer review by another layperson. Recorded amounts of time spent by each student in completing the answer sheet ranged from 5 to 42 hours, over periods of up to 5 separate days, with mean time spent of 9.46 (5.18) hours.

As with the initial study that involved FDEs, scoring for the Laypersons Group was based on the total of 52 possible selections from the pool of known writers, which meant that each of the 52 possible answers had a value of 1.923 points. A perfect score of 100 points would require 100% correct pairings of the 52 known documents with the 43 questioned documents. For each incorrect response, or answer left blank, 1.923 points were deducted. For reporting purposes, scores were converted to percent correct and rounded up or down to the nearest whole number.

As a group, the mean score for all participating laypersons was 76.43%% (18.12.) The highest score was 100% (scored by only 1 layperson), the lowest score was 33% (scored by 1 layperson), and the most frequent score was 90% (scored by 6 laypersons.) Remaining scores ran the gamut: 5 laypersons scored 98%, while another 5 scored 57%. Scores of 94%, 92%, 86%, 80% and 71% were attained by 3 laypersons in each of these percentiles; scores of 63% and 61% were attained by 2 laypersons each; and scores of 84%, 76%, 73%, 69%, 65%, 53%, 51%, 43%, and 37% were attained by 1 layperson each. Of the 4 laypersons who engaged in some form of peer review, **none** achieved a perfect score.

Association Between Time to Complete Matching Task and Accuracy Score

Overall, there was no correlation between the time spent on the task and the scores attained: r−0.13 for all 46 Laypersons. However, for male students only (n−9), there was a statistically significant relationship between higher scores and longer time spent on the task (r−0.77; p−0.015.) This was not true for females (n−37; r−0.11.) No difference between the genders was found in time spent on the task or scores attained. Females in the group took longer overall to complete the task: mean time for males of 8.56 (0.72) hours; and mean time for females of 9.67 (5.76) hours. However, the difference in scores was insignificant: mean scores for males−70.0 (22.7); and mean scores for females−78 (16.2.)

Table 1. Table depicts comparison of highest, lowest, most frequent, and average scores attained by Laypersons and FDEs who examined and compared the same homogenous writing specimens.

Scores	Forensic Document Examiners	Laypersons
Highest	100% (29)	100% (1)
Lowest	85% (1)	33% (1)
Mode	100%(29)	90% (6)
Group Average	98%	76%

Comparison of Scores by Laypersons to Scores by FDEs

The FDEs who examined the same documents in the initial study performed quite well when compared to the Laypersons Group [Table 1.] The overall average accuracy score for the FDE Group was 98%—22 percentage points higher than overall average score of 76% attained by the Laypersons Group. Furthermore, the scores within the FDE Group were higher in general than those attained in the Laypersons Group [Figure 1.] Twenty-nine FDEs scored 100%, 14 scored 95%, 2 scored 94%, 2 scored 90%, 1 scored 88%, and 1 scored 85%. By comparison, approximately 65% of the Laypersons Group scored at or below the single LOWEST score of 85% attained in the FDE Group. Using the conventional 95% accuracy as a cut-point, the proportion of FDEs (86%) was highly significantly greater than the proportion of laypersons (13%) meeting or exceeding this threshold of accuracy (χ2–51.03; $p<0.00001$.)

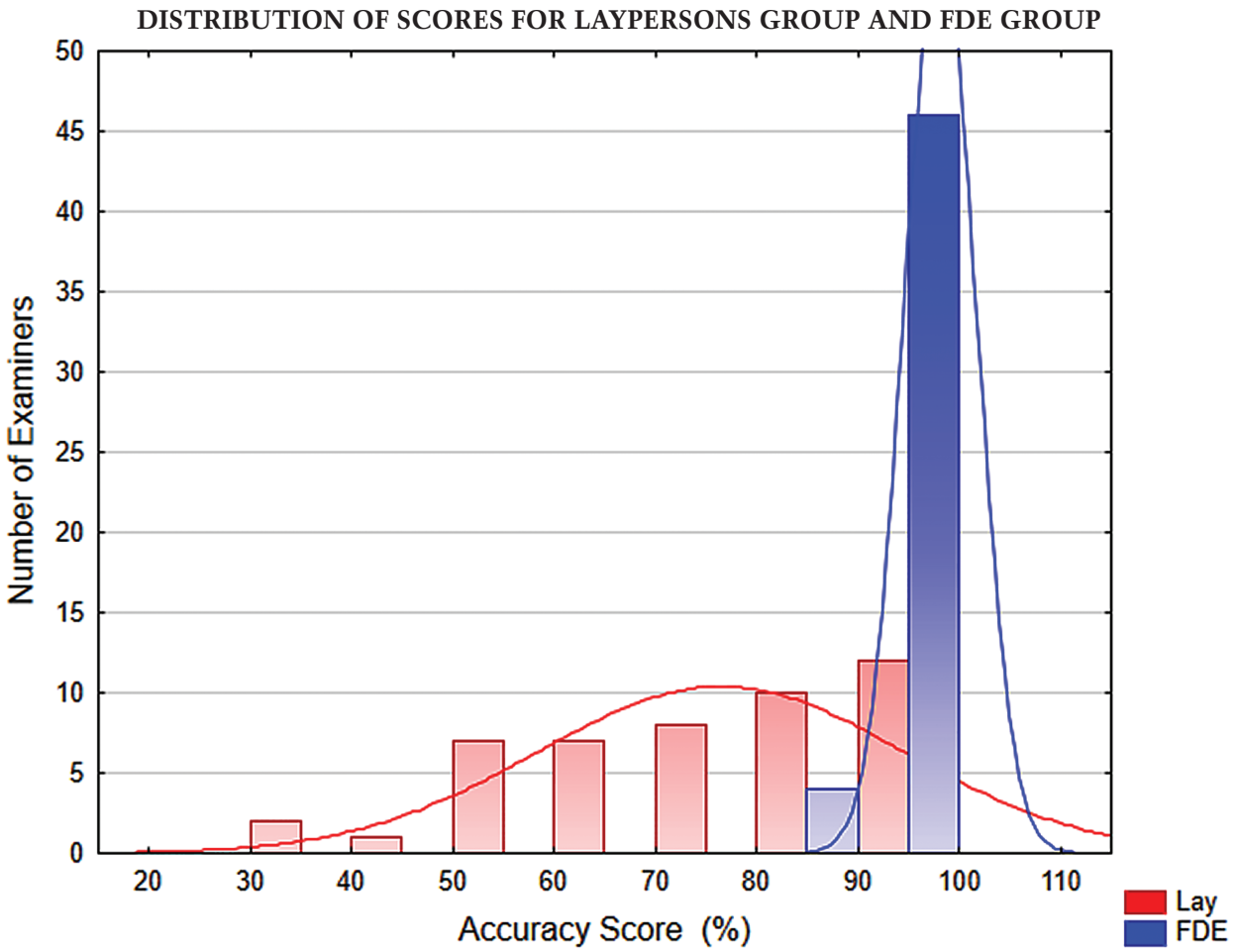

Fig.1. Figure 1 shows the distribution of correct scores for Laypersons and FDE subjects plotted as histograms.

A comparison was performed of the total number of incorrect associations made by both groups: those made by the FDE Group reported in 2009, and those made by Laypersons Group and now being reported. These incorrect associations were tallied by counting the number of times each Q and each K document was erroneously paired. Results showed that the FDE Group made a total of 67 incorrect associations, while the Laypersons Group made a total of 861 incorrect associations. [Tables 2 and 3 combined.]

Given the fact that every participant had exposure to the same group of 52 K specimens when attempting to make associations to each of the 43 Q documents, there was a total of 2,236 possible pairings that could have been made (43 x 52=2,236.) Therefore, a comparison of incorrect pairings made by both groups results in overall error rates of approximately 3% for the FDE Group (67/2,236=0.02996) and approximately 39% for the Laypersons Group (861/2,236=0.3850.)

Problematic Specimens for Laypersons and FDEs

In the group of 43 questioned documents, a total of 13 Qs (or 30% of all Qs) were found to be problematic for FDEs and resulted in at least one incorrect association. For the purpose of this research, we define problematic documents as those samples that failed to be correctly associated with their authors. Some of these incorrect associations occurred when the Q document contained fewer than 20 words. These incorrect associations were discussed at length in the published results of the initial study [23.] For the Laypersons Group in the study now being reported, **ALL** 43 of the questioned documents (100% of them) proved problematic, and were incorrectly associated one or more times. [Table 2.]

In the group of 52 known documents, a total of 11 Ks (or 21% of all Ks) were found to be problematic for FDEs in the initial study, and resulted in at least one incorrect association by members of the FDE Group. For the Laypersons Group in the study being reported here, 49 of the known documents (94% of all Ks) proved problematic, and were incorrectly associated one or more times. Only 3 of the Ks were not incorrectly associated by the Laypersons group. [Table 3.]

Table 2. Comparison of incorrect associations made of Q documents by FDEs and Laypersons. 13 of the Qs were found to be problematic for the FDE Group, however, all 43 Qs were found to be problematic for the Laypersons.

PROBLEMATIC QUESTIONED DOCUMENTS		
QUESTIONED DOCUMENT NUMBER	Number of Incorrect Associations by LAYPERSONS	Number of Incorrect Associations by FDEs
Q1	6	
Q2	3	
Q3	2	
Q4	7	
Q5	5	
Q6	1	
Q7	13	1
Q8	2	
Q9	22	
Q10	14	1
Q11	2	
Q12	3	
Q13	17	
Q14	3	
Q15	12	
Q16	21	1
Q17	10	
Q18	15	1
Q19	18	
Q20	9	1
Q21	4	
Q22	12	
Q23	9	
Q24	2	
Q25	23	5
Q26	1	
Q27	17	
Q28	17	8
Q29	6	
Q30	6	
Q31	19	2
Q32	24	11
Q33	2	
Q34	7	
Q35	10	1
Q36	12	
Q37	18	1
Q38	4	
Q39	34	6
Q40	9	
Q41	8	
Q42	13	
Q43	14	1
TOTAL NUMBER OF ERRONEOUS ASSOCIATIONS	456	40

Table 3. Comparison of incorrect associations of K documents made by FDEs and Laypersons. 11 of the Ks were found to be problematic for the FDE Group; however, all but 3 Ks were found to be problematic for the Laypersons.

PROBLEMATIC KNOWN DOCUMENTS		
KNOWN DOCUMENT NUMBER	Number of Incorrect Associations by LAYPERSONS	Number of Incorrect Associations by FDEs
K1	4	
K2	1	
K3	15	
K4	4	
K5	3	
K6		
K7	10	
K8	17	
K9	22	3
K10	13	
K11	1	
K12	4	
K13	5	
K14	5	
K15	3	
K16	15	1
K17	4	
K18	14	2
K19	7	
K20	14	
K21	7	
K22	5	
K23	12	
K24	5	1
K25	2	
K26	7	
K27	4	
K28	15	2
K29	2	
K30	4	
K31		
K32	33	11
K33	11	
K34	5	
K35	8	
K36	11	
K37		
K38	14	1
K39	15	1
K40	13	1
K41	1	
K42	6	
K43	3	
K44	5	
K45	2	
K46	3	
K47	15	
K48	10	3
K49	3	
K50	14	1
K51	6	
K52	3	
TOTAL NUMBER OF ERRONEOUS ASSOCIATIONS	405	27

Commentary by Laypersons

Several participants in the Laypersons Group commented on the answer sheets about their participation in the study. No comments were made about the constriction of not being permitted to render inconclusive or qualified opinions: a matter of great consternation to, and commentary by, many of the FDEs who participated in the initial study. The Laypersons Group, all of whom were students, seemed readily willing to accept that answers left blank and incorrect associations would be scored as "wrong" answers and result in deducted points.

Some in the Laypersons Group commented about their strategy for completing the task of examination and comparison. For example, one layperson referred to using "the process of elimination" by choosing to associate a writer with a particular specimen, then removing that writer from the pool of those remaining. Others, however, did not seem to have a particular strategy. Many laypersons chose to associate the same K writing specimen over and over again on the same answer sheet with several different Qs, perhaps in the hope that at least one selection would generate a correct answer. This is noteworthy, particularly given the admonishment by the author (Durina) to the Laypersons Group prior to commencing their task, reminding them to treat the specimens as actual forensic evidence that could determine guilt or innocence. It is also noteworthy that this strategy of associating multiple authors with the same Q document was never observed in the FDE Group, perhaps because one of the tenets of forensic document examination is once a questioned document is "identified" as

having been written by a particular writer, the same questioned document cannot be "identified" yet again as having been written by a different writer. Critics might pontificate that the behavior of laypersons identifying multiple authors for the same questioned document indicates either participant disinterest in the project, or a lackadaisical approach. However, given this Laypersons Group's interest in pursuing careers in forensic science, and given each participant's completion of at least 2 prior courses covering the importance of physical evidence and its role in the criminal justice system, neither scenario seems likely.

Some in the Laypersons Group remarked that it took several sittings over many days to complete the task of associating the Q documents with the K specimens. Yet, many also remarked that they found the process rather "easy" to complete.

For example, some of the comments in this vein included:

"Only the last 5 were hard to match." (This layperson scored 51%.)

"The samples were easily matched, but I had to look at the last few more closely." (This layperson scored 63%.)

"It seemed easy to pick them out. I want to know if it was as easy as I thought." (This layperson scored 73%.)

Much like results reported in a study conducted by Dr. Merlino, there appeared to be a higher degree of confidence among the Laypersons Group in making the correct associations than their accuracy levels warranted [25.] The commentaries of this nature by the laypersons are in stark contrast to those made by the FDEs in the initial study. Unlike participants in the Laypersons Group, many of the FDE participants commented that they found the same task of comparing the same group of homogeneous writings to be quite challenging. In fact, not a single FDE participant in the initial study implied in his/her comments that the assignment was performed with ease.

As all Laypersons Group participants were in the age brackets and classifications of persons subject to being called upon for jury duty, the nature of the comments made concerning the ease of determining authorship is somewhat troubling. One cannot help but wonder how many laypersons view examination and comparison of forensic evidence as a simple task they could easily perform, and what the impact of these mindsets might be in courts of law where such mindsets might affect the presentation, acceptance, and weight of forensic evidence such as handwriting.

Limitations

As is the case with most research studies, this one was subject to certain limitations. The limitations of the writing samples examined were the same as those reported in the initial study: all consisted of photocopies and digital images. Writers were not observed when preparing the writing specimens, and no data regarding the physical limitations or mental conditions of the writers at the time of the exercise was obtained. Additional limitations within the body of specimens included the fact that the Qs varied in length and were not directly comparable to the content of the Ks. Differences in writing styles were also evident in that some of the Q documents were written in a less formal style than the K documents, where the writers appeared to make more of an attempt to adhere to the Palmer copybook style.

In an attempt to replicate the conditions and constraints for completing the answer sheet imposed upon the FDEs who participated in the initial study, the Laypersons participating in this study were not permitted to render qualified opinions. As in the initial study involving the FDEs, inconclusive opinions and answers left blank by the Laypersons were marked as incorrect.

Laypersons were given over one month to complete their answer sheets and were not observed performing the comparison tasks. Some stated it took them many hours over several days to complete their answer sheets, and many expressed a great interest in, and gratitude for, the opportunity to be given a "hands on" opportunity to examine and compare forensic document evidence. However, not all students provided commentary. It is unknown how some of the students went about the comparison process, or how much time was spent by each student on each particular writing specimen.

Other than age, gender, ability to read cursive writing, and lack of prior scientific training or experience in forensic document examination, no background information about the Laypersons was obtained (i.e. individual health issues, scholastic abilities, and conditions such as form/color blindness, are unknown.) The only incentive given to the participating Laypersons, all of whom were students, was an opportunity to gain extra credit points toward their final grade at the

end of the semester for correct answers rendered. Although many expressed a desire for these extra credit points, it is unknown how strong an incentive this was to each individual participant.

Conclusion

There was a significant difference between the FDEs' accuracy and that of the Laypersons Group. The group of FDEs was able to determine authorship correctly with an average accuracy score of 98%, while the Laypersons Group's average accuracy score was only 76%. Overall error rates in these two studies were determined to be approximately 3% for the FDEs Group, compared to approximately 39% for the Laypersons Group. Peer review increased the accuracy rate to 100% for all 15 FDEs who engaged in peer review, but did NOT improve accuracy for the 4 laypersons who engaged in it. Samples that were problematic for FDEs were limited to 13 Q documents and 11 K documents, yet all 43 Q documents and 49 of the 52 K documents were found to be problematic for the Laypersons Group.

Prior research in this area by others has shown that expertise of professional FDEs is superior to that of laypersons performing the same tasks. While these results vary from study to study, as do the comparison tasks the participants have been asked to perform (e.g. comparisons of signatures only, writings that may be simulated or disguised, hand printing, and numerals), accuracy levels of FDEs have been found to be consistently higher than those of laypersons. This particular study involved only samples of cursive writings that were naturally written by a group of writers who met the recommendations noted in a critical commentary in 2003 which stated: "A study of handwriting would be far more convincing if the writers in the sample had all grown up in the same neighborhood, gone to the same school and had been taught by the same teachers" [26.] The writings examined and compared in this study from that homogeneous group of writers were not simulated or disguised, however, driving forces for writing variation were low, and abundant similar class characteristics throughout the writing specimens presented challenges to the determination of authorship. It is hoped that the results reported here will provide additional supporting empirical evidence that professional expertise in handwriting examination **does** exist, and that professional forensic document examiners demonstrate greater skill in the examination and comparison of handwriting than that demonstrated by laypersons.

Acknowledgements

The authors express their deepest gratitude to the 46 Grossmont College Students who participated in this study as the Laypersons Group and diligently reported their results; to Ms. Tina Young, Professor at Grossmont College, for generously allowing author Durina access to the College and its students and for offering incentive for student participation in the study; to the former students of Our Lady Perpetual Help Elementary School in Brooklyn, NY and the Sisters of St. Joseph who taught there, for the donation of their handwritten specimens used in these studies; and to the 49 Forensic Document Examiners from Australia, Canada, Finland, the Netherlands, and the United States who participated in the initial study in 2008-2009.

References

1. U.S. Supreme Court ruling, *Daubert v. Merrell Dow Pharmaceuticals, Inc*. 509 US 579, 1993.

2. U.S. Supreme Court ruling, *United States v. Starzecpyzel,* 880 F. Supp. 1027 (S.D.N.Y.) 1995.

3. U.S. Supreme Court ruling, *General Electric Co. v. Joiner*, 76 F. 3d., 524, 1997.

4. U.S. Supreme Court ruling, *Kumho Tire Co. v. Patrick Carmichael*, 131 F 3d 1443, 1999.

5. Saks, M., Denbeaux, M., and Risinger, M. (1989) Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification "Expertise", *Pennsylvania Law Review*, Vol. 137, No. 3, pp. 731-792.

6. Risinger, D. M., and Saks, M. (1996) Science and NonScience in the Courts: Daubert Meets Handwriting Identification Expertise, *Iowa Law Review*, Vol. 82, pp. 21-74.

7. 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.

8. Kam, M., Fielding, G., and Conn, R. (1997)

Writer Identification by Professional Document Examiners, *Journal of Forensic Sciences*, Vol. 42, pp. 778-786.

9. 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. 1000-1004.

10. Kam, M., Gummadidala, K., Fielding, G., and Conn, R. (2001) Signature Authentication by Forensic Document Examiners, *Journal of Forensic Sciences*, Vol. 46, pp. 884-888.

11. Found, B., Sita J., Rogers D. (1999) The development of a program for characterising forensic handwriting examiners' expertise: signature examination pilot study, *Journal of Forensic Document Examination*, Vol. 12, pp. 69-80.

12. Sita, J., Found, B., and Rogers D.K. (2002) Forensic Handwriting Examiners' Expertise for Signature Comparison, *Journal of Forensic Sciences*, Vol. 47, No. 5; pp. 1117-1124.

13. Bird, C., Found, B., and Rogers D.K. (2010) Forensic Document Examiners' Skill in Distinguishing Between Natural and Disguised Handwriting Behaviors, *Journal of Forensic Sciences*, Vol.55, No. 5; pp. 1291-1295.

14. Strengthening Forensic Science in the United States: A Path Forward, Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council, ISBN: 0-309-13131-6, (2009), PDF available electronically from *National Academies Press* at: http://www.nap.edu/catalog/12589.html

15. Ibid.

16. Scientific Working Group for Document Examination, Published Standards, Available from *SWGDOC.org* at http://www.SWGDOC.org/index.php/Standards/published-standards.

17. Vargas, R. (2008) Standards for Questioned Document Examination: Preparing Document Examiners for the Judicial System of the 21st Century, *ASTM Standardization News*, November/December 2008, pp. 51-55.

18. Merlino, M. (2015) Validity, Reliability, Accuracy, and Bias in Forensic Signature Identification, Document No. 248565, pp. 1-1470, PDF available electronically from *NCJRS* at http://www.NCJRS.gov/app/publication/abstract/aspx?ID#270668

19. Kam, M., Abichandani, P., and Hewett, T (2015) Simulation Detection in Handwritten Documents by Forensic Document Examiners, *Journal of Forensic Sciences*, Vol. 60, pp. 936-941.

20. Srihari, S., Cha, S., Arora, H. and Lee, S. (2002) Individuality of Handwriting, *Journal of Forensic Science*, Vol. 47, No. 4; pp. 856-872.

21. Durina, M. and Caligiuri, M. (2009) The Determination of Authorship from a Homogenous Group of Writers, *Journal of the American Society of Questioned Document Examiners, Inc.*, Vol. 12, No. 2, pp. 77-89.

22. Ibid.

23. Ibid.

24. Ibid.

25. Merlino, M. (2015) Validity, Reliability, Accuracy, and Bias in Forensic Signature Identification, Document No. 248565, p. 1424, PDF available electronically from NCJRS at http://www.NCJRS.gov/app/publication/abstract/aspx?ID#270668

26. Saks (2003) Commentary on: Srihari, S., Cha, S., Arora, H. and Lee, S. (2002) Individuality of Handwriting, *Journal of Forensic Science*, 47(4), 856-872.