

Online Proficiency Testing for Signature Comparison by Forensic Document Examiners and Non-Examiners

Lindsay Holmes, Brent Ostrum, and Andrew Barton

Forensic Document Examiners, Canada Border Services Agency

14 Colonnade Road, Suite 280, Ottawa, Ontario, K2E 7M6

The present study was conducted to develop and test a new online method of proficiency testing for the completion of a signature comparison trail. The web-based test consisted of 14 specimen signatures and 50 questioned signatures with the latter consisting of 8 'genuine' signatures and 42 'simulated' signatures. Opinions regarding the authorship of each questioned signature were submitted by a group of 32 Forensic Document Examiners (FDEs) as well as a group of 32 non-examiners. Results from the online test for FDE participants were compared with the results from a paper-based signature comparison trial conducted by La Trobe University Forensic Expertise Profiling Laboratory that had used signatures from the same specimen writer and simulators. This comparison revealed that there was no significant difference between the FDEs who participated in the online study and the FDEs from the La Trobe study. In addition, FDE participants performed better than the non-examiner group, providing more correct opinions and fewer misleading opinions. The results of the study provide further validation of the existence of an expertise in the area of signature identification. Support is also provided for the use of online proficiency testing as an alternative to the traditional paper method.

Introduction

Forensic Document Examiners (FDEs) examine documents for many purposes such as assessing possible authorship, determining how a document was created, or evaluating whether a document has been altered or is fraudulent. One common type of examination is the determination of authorship of a written document by conducting a handwriting or signature comparison (Kelly and Lindblom 2006). This type of examination involves the comparison of questioned writings with specimen samples to determine if the two were produced by the same person. After examining and comparing the samples, the FDE may give an opinion regarding the authorship of the questioned writing. The issue of whether the FDE's opinion should be given more weight than the opinion of a layperson or non-examiner has come up in the scientific literature (Risinger, Denbeaux, and Saks 1989; Galbraith, Galbraith, and Galbraith 1995). To address this issue, proficiency tests have been developed to establish whether FDEs possess an expertise in the areas of handwriting and signature identification (Kam,

Wetstein, and Conn 1994; Kam, Fielding, and Conn 1997; Kam, Gummadidala, Fielding, and Conn 2001; Sita, Found, and Rogers 2002).

In general, proficiency tests present the examiner with tasks similar to those they encounter during their everyday casework but provide the benefit of knowing the correct answer for the test samples. These tests are often used as a form of training for FDEs but have also been used to determine if forensic handwriting and signature identification should be considered an expertise (Kam, Wetstein, and Conn 1994; Kam, Fielding, and Conn 1997; Kam, Gummadidala, Fielding, and Conn 2001; Sita, Found, and Rogers 2002). This is done by comparing the FDEs performance on the proficiency test to the performance of non-examiners, who have no experience in document examination. Several studies, particularly those that include handwriting and signature examinations, have been able to demonstrate that the abilities of professional FDEs are superior to the abilities of the average person (Kam, Wetstein, and Conn 1994; Kam, Fielding, and Conn 1997;

Kam, Gummadidala, Fielding, and Conn 2001; Sita, Found, and Rogers 2002). Although it has been shown that expertise exists, validation studies should be continued in order to further validate those findings. Additional studies are also needed in order to clarify the specific nature of any differences between FDEs and non-examiners and to extend the testing into different areas of document examination.

The format of the proficiency tests conducted to date by the Forensic Expertise Profiling Laboratory (FEPL) at La Trobe University in Melbourne, Australia (Found and Rogers 2006(1); Found and Rogers 2006(3); Found and Rogers 2007(1); Found and Rogers 2007(2); Found and Rogers 2008(1); Found and Rogers 2008(2)) and by Skill-Task, Training Assessment and Research, Inc. (ST2AR) in the United States (Hammond and Found 2008; Hammond and Found 2009) involving the examination of handwriting and signatures, have all been reasonably consistent. First, packages containing all of the testing material were distributed to the participants. Participants have included FDEs with varying levels of experience from different government and/or private agencies. Researchers have also had the option to include a group of non-examiner participants in order to compare their abilities with those of the FDEs. Depending on the specific study, packages have contained a variety of testing materials. A typical handwriting or signature examination has included three main elements; the specimen samples, the questioned samples, and an opinion scale.

During the test, participants were requested to compare the set of questioned samples with the specimen set and give an opinion regarding the authorship of each questioned sample using a specific scale. The scale of opinions normally ranged from concluding the questioned sample was written by the specimen writer to concluding that the questioned sample was not written by the specimen writer with a total of five possible opinions.

Test results for all questioned samples would then be submitted to the researcher for analysis. The analysis of the results consisted of calculating the rate of correct, misleading and inconclusive opinions made by each participant as well as the opinions given for each type of questioned signature (genuine, simulated, disguised, spurious, etc). Analyzing the participants test responses and reporting the findings concluded the experimental process.

This method has proven to be successful in establishing the presence of an expertise in the area of forensic handwriting identification. While successful in that respect, it does involve a very lengthy and often expensive process. Many copies of the test package must be constructed and sent to all participants, some of whom are located in different countries. The original samples produced by the specimen writer and simulators would not generally be sent out to participants. Instead, high quality copies (photographs or prints) of the samples are made. The samples have also been scanned and sent to the participants on a CD. Upon completion of the test by a participant, the answer sheet must be mailed or faxed back to the researcher for collation and analysis. This may involve recording the responses from a few to several hundred participants followed by the inputting of data into a computer program to perform the statistical analysis. In addition to the time and cost factors, there is also the concern about possible transcription or data entry errors.

The present study tested an online proficiency test as an alternative to the paper version described above. The online test, which involved a signature comparison, followed the same general format as the previously conducted tests except that test packages containing printed copies of the samples were not provided to the participants. Instead, all specimen and questioned material were made available to participants through an online website. On the website, each test page presented one questioned signature and a set of specimen signatures. After making their comparison using the images on the website, participants entered an opinion regarding the authorship of each questioned signature onto the webpage and submitted their response to the researcher. By having the test online, the time and cost involved in making copies of each sample and mailing them to the participants was completely eliminated. In the present study the participant's test responses were received through email and were then manually entered into a database for analysis. Although manual data entry is not ideal and may lead to data errors, this approach was taken to minimize the complexity of the design process for the website. It would be possible to design a website to automatically input the participant's responses into a database on the researcher's computer and have the statistical analysis done as responses are received. This would not only decrease the time needed to

complete the statistical analysis but would also minimize the amount of transcription error that may occur while inputting data and will be done for any future studies.

The set of specimen and questioned signatures included in the online study were borrowed from the FEPL at La Trobe University. This laboratory has completed numerous studies in the area of forensic document examination including several signature comparison studies. La Trobe conducted a signature comparison study in 2007 (Found and Rogers 2006(2); Found and Rogers 2007(2)) that involved signatures from the same 'master' set as those used in the current study. Participants in both tests were also given similar five point opinion scales to use when making their conclusions about each signature. The similarities between the two test formats allowed the results to be directly compared. It was believed that this comparison would reveal whether the testing method (paper-based or online) affected the opinions given by FDEs in any significant manner.

The La Trobe study found that out of the 2500 opinions given by FDEs on 100 questioned signatures, 94.6% were correct, 1.8% were misleading, and 3.6% were inconclusive. It was expected that the results from the current study would not be significantly different from these. The two test formats were also compared in terms of the called opinion rates (rate of correct and misleading opinions when inconclusive opinions are removed), the strength of the opinions given (qualified or unqualified), and the opinion rates for each type of questioned signature (genuine and simulated).

The experience of each FDE with proficiency testing was also considered. Since the signatures used in the present study were taken from a signature set used in the 2007 La Trobe study there is a possibility that some FDEs may have been participants in both. FDEs who indicated that they had previously participated in La Trobe proficiency tests were compared to the FDEs who did not have previous experience to determine if experience in the La Trobe testing affected their performance on the online test.

It was believed that obtaining similar results using two test formats would indicate that online proficiency testing could be an alternative to the traditional paper version. However, if significant differences were seen in the results it would lead to an examination of why FDE performance changed when similar information was presented using a different testing format.

Unlike the 2007 La Trobe study, the present study included a group of participants with no experience in forensic document examination to be compared to the FDE group. As a result, the responses provided by non-examiners were only compared to the results of the FDEs in the current study. Previous studies have found that FDEs performed better on proficiency tests involving handwriting and signature comparisons than non-examiners (Kam, Wetstein, and Conn 1994; Kam, Fielding, and Conn 1997; Kam, Gummadidala, Fielding, and Conn 2001; Sita, Found, and Rogers 2002). However, not all studies found that the opinions given by document examiners were significantly different from those of non-examiners in all areas. Studies have shown that the opinions of document examiners and non-examiners are similar when it comes to making correct matches, but differ in regards to misleading and inconclusive opinions (Kam, Fielding, and Conn 1997; Sita, Found, and Rogers 2002). A signature comparison study by Sita et al. (2002) found no difference in the number of correct opinions given by the FDE and non-examiner groups. However, FDEs gave significantly fewer misleading opinions and significantly more inconclusive opinions. The difference in the number of inconclusive opinions may represent a level of conservatism by the FDEs and could be a reflection of how they may proceed in their casework since concluding that a signature was written by a certain author when it was not can have considerable negative consequences.

Along with a comparison of the overall results, the called opinions of the FDE and non-examiner groups were also compared. Lastly, a comparison was made between the performance of the FDE and non-examiner groups with respect to genuine questioned signatures versus simulated questioned signatures. This comparison may give some insight into the areas where FDEs and non-examiners differ in their ability to give correct authorship opinions. The presence of significant differences between the FDE group and the non-examiner group would give further support to previous finding that FDEs have a legitimate expertise in signature identification.

Methods and Materials

In this study, forensic document examiners and non-examiners were asked to determine if 50 questioned signatures were genuine or simulated using an online proficiency test. Before the

official study began, pilot testing was conducted to determine if the testing website was working correctly. Following the pilot tests, minor changes were made to the online test to improve functionality, such as including lower resolution images on the test pages with the option to view a higher resolution image in a separate window.

Participants

Two groups of participants were involved in this study, forensic document examiners and non-examiners. The FDE group consisted of 32 practicing document examiners. Participants in this group were from several different countries including six from the USA, four from Canada, four from Australia and New Zealand, and 18 from European countries. The amount of experience with proficiency testing as well as the type of tests, if any, that were participated in, varied amongst the FDE group. FDE participants were all contacted through email, requesting their participation in the study. If the contacted individual was willing to participate they were directed to an internet link that would open the online test. Approximately 230 FDEs were invited to participate in the signature test. Out of those invited, 32 completed the test giving a final response rate of 13.9%.

The second group of participants was called the non-examiner group and consisted of 32 undergraduate students from Trent University in Peterborough, Ontario. The age and area of study varied among the non-examiner participants. Twenty-six of the 32 participants in this group indicated that they had no previous knowledge or experience with authorship determination. There were, however, six non-examiner participants that indicated they had limited knowledge or experience. Many of the non-examiner participants are members of the forensic science undergraduate program at Trent University. Though no course material directly discussed forensic document examination or signature identification, students may have felt that their forensic science background provided them with some experience. A guest lecture was also given by an FDE to students at Trent University. It was not a mandatory lecture though several students attended. If non-examiner participants had attended this lecture, they may have felt that they gained some experience in signature identification. Non-examiner participants were contacted through email as well as in-class announcements

to request participation. Approximately 200 Trent students were invited to participate in the online test. Out of those invited, 32 completed the test giving a final response rate of 16%.

In both groups, there were participants who submitted their registration information but did not begin the test itself, or began the test but failed to complete it. In total, 95 individuals submitted a registration but only 64 completed the signature comparison giving a drop-out rate of 32.6%.

For both groups, participation in the study was completely voluntary and no rewards were given to those who chose to participate.

Signature Samples

The questioned and specimen signatures used in the online proficiency test were borrowed from The Forensic Expertise Profiling Laboratory at La Trobe University in Melbourne, Australia.

The signatures that were borrowed included a set of 40 genuine signatures written by a specimen writer. The adult writer was instructed to sign her natural signature nine times a day for five days using a blue ballpoint pen (Found and Rogers 2006(2)). Fourteen of the signatures had been used as specimen signatures in the 2007 La Trobe study and were used again as the specimen signatures in the present study. Eight of the remaining signatures were removed from the set as they had been used in the previous study as part of the questioned signature set. A sequence generator from www.random.org was used to put the remaining 18 signatures into a random order. The first eight signatures on this list were chosen to be incorporated into the questioned signature set.

Also included in the signature package was a set of 42 simulated signatures. Twenty-four writers were instructed to create 20 simulations of the specimen signature using a blue ballpoint pen. Each simulator was given a set of three specimen signatures to use as a reference and was instructed to produce simulations without tracing the specimens. One of these writers (writer 11) was instructed to write 20 signatures a day for a period of ten days to give a total of 200 simulated signatures (Found and Rogers 2006(2)). One hundred of the simulated signatures had been used in the previous La Trobe study and therefore had to be excluded from the questioned signature set of the current study. The complete set of simulated signatures was placed in a random order using the

sequence generator at www.random.org. From the random list of signatures that was generated, the first 42 signatures that had not been used in the previous study were chosen to be a part of the questioned signature set.

The sequence generator was used a third time to randomize the order of appearance of the 50 signatures included in the questioned signature set. Though a random sequence of the 50 questioned signatures was generated, the same sequence was viewed by each participant. Once the specimen and questioned signatures had been selected they were scanned at a resolution of 600dpi in 24-bit colour using an Epson Perfection 4990 Photo scanner. Images were saved as high-quality JPEG images for use on the website.

It should be noted that due to the availability of signature samples it was not possible to add disguised signatures to the questioned signature set. During a signature examination an FDE must consider all possibilities for the questioned signature including the possibility that the features of the writing were intentionally altered by the specimen writing in an attempt to conceal his or her identity or to later deny authorship. The absence of disguised signatures reduced the direct comparability of the test to casework; however, it was still possible to make a comparison between FDEs and non-examiners as well as between the online test and previous test.

Test Procedure

A website was created through which the participants would complete the proficiency test. The website was posted on MyWebSpace which was provided by Trent University and was available at www.people.trentu.ca/lindsayholmes. The online test included a letter of information, an instruction page, a registration page, 50 test pages and a debriefing page.

On the registration page, participants were asked to provide personal and professional information. Personal questions included the participant's name, e-mail address and country of residence. Professional questions asked whether the participant is a professional document examiner and if so, if they are certified and by what certifying body, their familiarity with authorship determinations and any previous experience with proficiency testing. After answering all of the mandatory questions, participants were given a chance to write down any additional comments.

The signature comparison test consisted of 50

test pages. Every page contained one questioned signature and 14 specimen signatures. Each of the signatures on the test page could be enlarged by clicking on the image. Participants were instructed to compare the questioned signature to the specimen signatures and score the questioned signature using a five point scale that was modeled after the one used in the 2007 La Trobe trial. The scale ranged from 1 to 5 and represented the following opinions:

1. The questioned signature was written by the specimen writer.
2. It is likely that the questioned signature was written by the specimen writer.
3. It is inconclusive whether the questioned signature was or was not written by the specimen writer.
4. It is likely that the questioned signature was not written by the specimen writer.
5. The questioned signature was not written by the specimen writer.

After giving an opinion for the questioned signature the participant would submit their answer to the researcher. After submitting their answer, the participant was directed to the next test page. Participants were required to submit an opinion for the questioned signature they were examining before being able to proceed to the next question.

After all 50 test pages had been completed the participants were directed to a debriefing page. This page allowed for comments about the test to be entered and sent to the researcher.

Test responses were separated from identifying information such as the participants name and email address upon being received by the researcher. Responses were sorted as belonging to the FDE group or the non-examiner group.

Data Analysis

The number of correct, misleading and inconclusive opinions was recorded for each individual participant. Using these results, the percentage of correct, misleading, inconclusive, correct called and inconclusive called opinions was calculated for each individual. The percentage of correct, misleading, and inconclusive opinions given for genuine questioned signatures and simulated questioned signatures was also calculated for each participant. Finally, the average percentage of opinions given at each opinions level was calculated for FDE participants only.

Two main comparisons were made in this study, one between the FDEs who participated in

Opinion	F-value	P-value	Significantly Different?
Percentage Correct	1.019	0.317185	No
Percentage Misleading	0.001874	0.965631	No
Percentage Inconclusive	1.126218	0.293221	No

Table 1 Trent vs. La Trobe ANOVA results for correct, misleading, and inconclusive opinions, $\alpha = 0.05$

	# Correct	# Misleading	# Inconclusive	% Correct	% Misleading	% Inconclusive	% Correct Called	% Misleading Called
FDEs	1455	28	117	90.94	1.75	7.14	98.11	1.88
Non- Examiners	1187	231	182	74.19	14.44	11.38	83.71	16.29

Table 2 Summary of correct, misleading and inconclusive opinions for FDEs and non-examiners

the online test and the FDEs who participated in the 2007 La Trobe paper based signature trial and a second between the FDEs and non-examiner participants who took part in the online proficiency test. For both of these comparisons a series of Analysis-of-Variance (ANOVA) tests were performed using WebStatistica Software provided by StatSoft to determine if we should reject the null hypothesis (H_0) that there will be no difference between the opinions given by the two groups of participants. All ANOVAs were performed using an alpha level of 0.05. The tests were used to compare the results with regards to the overall number of correct, misleading, and inconclusive opinions, the average number of correct called opinions and misleading called opinions, the number of correct, misleading, and inconclusive opinions given for each questioned signature type (genuine and simulated), and the number of opinions given at each opinion level. ANOVAs were also performed to determine if the FDE and non-examiner groups of participants performed better with genuine or simulated questioned signatures.

Ethics Approval

Approval for this study was obtained from the Trent University Research Ethics Board. Approval was granted on the basis that participants were informed of their responsibilities before beginning the examination. It was also decided that the individual results of the participants would remain confidential and no identifying information would be included in the report of the study.

Results

Trent Study versus 2007 La Trobe Study (FDEs only)

The FDE group in the present study was compared to the FDE group from the 2007 signature comparison study by the La Trobe University Forensic Expertise Profiling Laboratory (Found and Rogers 2007[2]). The 32 FDE participants from the current study gave a total of 1600 opinions on 50 questioned signatures. Out of these opinions 1455 (90.94%) were correct, 28 (1.75%) were misleading, and 117 (7.31%) were inconclusive. The 25 FDEs from the 2007 La Trobe study gave a total of 2500 opinions on 100 questioned signatures. Out of these opinions 2365 (94.6%) were correct, 45 (1.8%) were misleading, and 90 (3.6%) were inconclusive.

A series of Main Effect ANOVAs was performed to determine if there was a significant difference between the opinions given by the two groups of participants. Table 1 displays the results of this analysis which showed no significant difference between the two groups of FDEs for the percentage of correct opinions ($F = 1.019$, $p = 0.317185$), misleading opinions ($F = 0.001874$, $p = 0.965631$), or inconclusive opinions ($F = 1.126218$, $p = 0.293221$).

A called opinion refers to any conclusion other than inconclusive (Sita, Found, and Rogers 2002). The number of called opinions was determined for each group of participants by removing inconclusive opinions entirely and considering only the correct and misleading opinions. For the Trent FDE group, the called rate of correct opinions was 98.11% and the called rate of misleading opinions was 1.88%. These results were not significantly different from rates found in the earlier La Trobe

Opinion	F-value	P-Value	Significantly Different?
Correct	19.486	0.000041	Yes
Misleading	45.31514	0.0	Yes
Inconclusive	1.48614	0.227434	No

Table 3 FDE vs. Non-examiner ANOVA results for correct, misleading, and inconclusive opinions, $\alpha = 0.05$

Signature Type and Opinion	Number of Opinions Given	Percentage of Opinions Given
Genuine Correct	240	93.75%
Genuine Misleading	7	2.73%
Genuine Inconclusive	9	3.52%
Simulated Correct	1215	90.40%
Simulated Misleading	21	1.56%
Simulated Inconclusive	108	8.04%

Table 4 FDE opinions for genuine and simulated questioned signatures

Opinion	F-value	P-value	Significantly Different?
Correct	0.764	0.385487	No
Misleading	0.98173	0.325624	No
Inconclusive	1.488867	0.227013	No

Table 5 FDE ANOVA results for genuine vs. simulated questioned signatures, $\alpha = 0.05$

study where the correct called rate was 98.13% ($F = 0$, $p = 0.948384$) and the misleading called rate was 1.9% ($F = 0.004229$, $p = 0.948384$). In fact, the results were almost identical in both studies.

Forensic Document Examiners versus Non-examiners (Trent Study participants)

FDE and non-examiner participants were asked to give an opinion as to whether each of 50 questioned signatures was or was not written by the specimen writer. Table 2 provides a summary of participant responses. The FDE participants gave a total of 1600 opinions. As noted earlier, out of these opinions 1455 (90.94%) were correct, 28 (1.75%) were misleading, and 117 (7.31%) were inconclusive. The non-examiner participants also gave a total of 1600 opinions. Out of these opinions, 1187 (74.19%) were correct, 231 (14.44%) were misleading, and 182 (11.38%) were inconclusive.

A series of Main Effect ANOVAs showed that FDEs gave significantly more correct opinions ($F = 19.486$, $p = 0.000041$) and significantly fewer misleading opinions than the non-examiner participants ($F = 45.31514$, $p = 0.0000$). There was no significant difference between the number of inconclusive opinions given by the FDEs and non-examiners ($F = 1.48614$, $p = 0.227434$) (see

Table 3).

When the called opinions were compared between FDEs and non-examiners, a significant difference was found to exist for both the correct called rates ($F = 47.711$, $p = 0.0000$) and the misleading called rates ($F = 47.71149$, $p = 0.0000$) with the performance of the FDE group being more accurate than the non-examiners in both situations.

Analysis by signature type

The set of 50 questioned signatures included in the online comparison test contained 42 simulated signatures and 8 genuine signatures. Trent FDEs expressed 93.75% correct opinions, 2.73% misleading opinions, and 3.52% inconclusive opinions on the genuine questioned signatures while simulated questioned signatures attracted 90.40% correct opinions, 1.56% misleading opinions, and 8.04% inconclusive opinions (see Table 4). There was no significant difference between the percentage of correct ($F = 0.764$, $p = 0.385487$), misleading ($F = 0.98173$, $p = 0.325624$), or inconclusive ($F = 1.488867$, $p = 0.227013$) responses given by Trent FDEs for genuine and simulated questioned signatures (see Table 5). These results indicate that the FDEs opinions were not significantly affected by the type of

Signature Type and Opinion	F-value	P-value	Significantly Different?
Genuine Correct	0.883	0.351588	No
Genuine Misleading	0.425175	0.517082	No
Genuine Inconclusive	0.366262	0.547538	No
Simulated Correct	0.852	0.360041	No
Simulated Misleading	0.042913	0.836654	No
Simulated Inconclusive	1.070936	0.305263	No

Table 6 Trent vs. La Trobe ANOVA results for genuine and simulated questioned signatures, $\alpha = 0.05$

Signature Type and Opinion	Number of Opinions Given	Percentage of Opinions Given
Genuine Correct	223	87.11%
Genuine Misleading	13	5.08%
Genuine Inconclusive	20	7.81%
Simulated Correct	964	71.73%
Simulated Misleading	218	16.22%
Simulated Inconclusive	162	12.05%

Table 7 Non-examiner opinions for genuine and simulated questioned signature

Opinion	F-value	P-value	Significantly Different?
Correct	12.378	0.000819	Yes
Misleading	16.45128	0.000142	Yes
Inconclusive	1.63333	0.206007	No

Table 8 Non-examiner ANOVA results for genuine vs. simulated questioned signatures, $\alpha = 0.05$

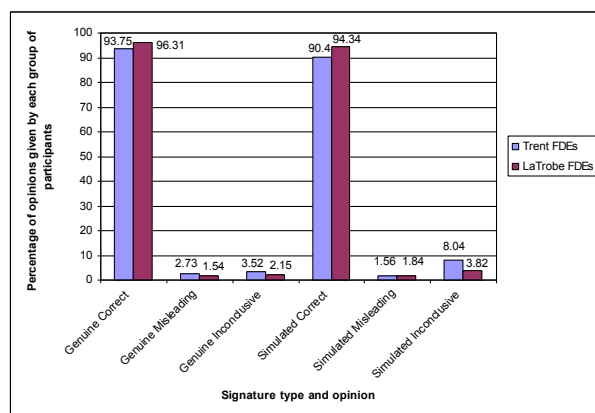


Figure 1 Comparison of Trent and La Trobe FDE results for genuine and simulated questioned signatures

questioned signature.

The 2007 La Trobe trial also included genuine and simulated questioned signatures. Figure 1 shows the percentage of opinions given by each group of FDEs for both signature types. Again, there was no significant difference found between

the opinions given by Trent FDEs and La Trobe FDEs for both genuine and simulated questioned signatures (see Table 6).

The non-examiner participants in the present study gave 87.11% correct opinions, 5.08% misleading opinions, and 7.81% inconclusive opinions on genuine questioned signatures and 71.73% correct opinions, 16.22% misleading opinions, and 12.05% inconclusive opinions on simulated questioned signatures (see Table 7). Non-examiners performed significantly better with genuine signatures, with more correct opinions ($F = 12.378$, $p = 0.000819$) and fewer misleading opinions ($F = 16.45128$, $p = 0.000142$). There was no difference between the percentage of inconclusive opinions given for genuine and simulated signatures ($F = 1.63333$, $p = 0.206007$) (see Table 8). Unlike the FDE participants, the non-examiners' opinions were affected by the type of questioned signature.

A comparison of the non-examiner and FDE groups showed that there was a significant difference in the percentage of correct opinions ($F = 17.637$, $p = 0.000087$) and misleading opinions

Signature Type and Opinion	F-value	P-value	Significantly Different?
Genuine Correct	3.147	0.080985	No
Genuine Misleading	1.50404	0.224688	No
Genuine Inconclusive	1.77688	0.187410	No
Simulated Correct	17.637	0.000087	Yes
Simulated Misleading	40.59111	0.000000	Yes
Simulated Inconclusive	1.12618	0.292707	No

Table 9 FDE vs. non-examiner ANOVA results for genuine and simulated questioned signatures, $\alpha = 0.05$

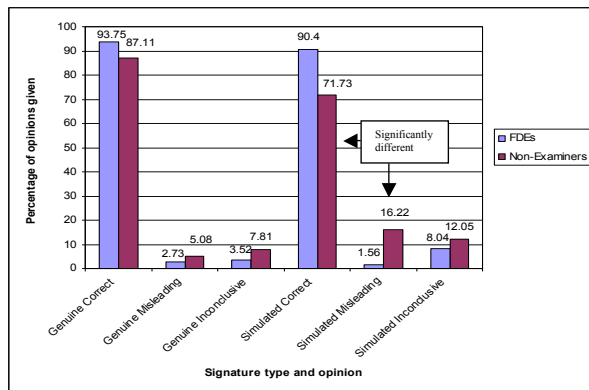


Figure 2 Comparison of FDE and non-examiner results for genuine and simulated questioned signatures

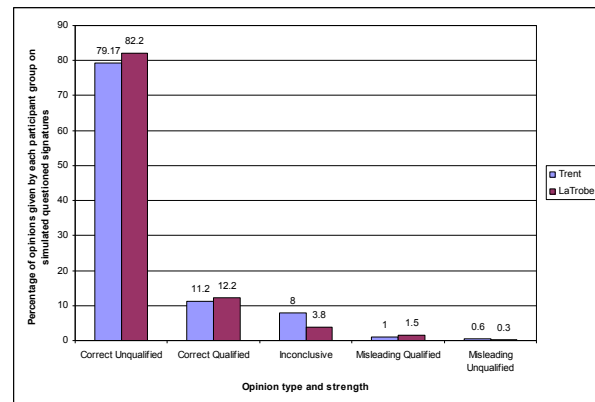


Figure 4 Comparison of Trent and La Trobe FDEs by opinion level for simulated questioned signatures.

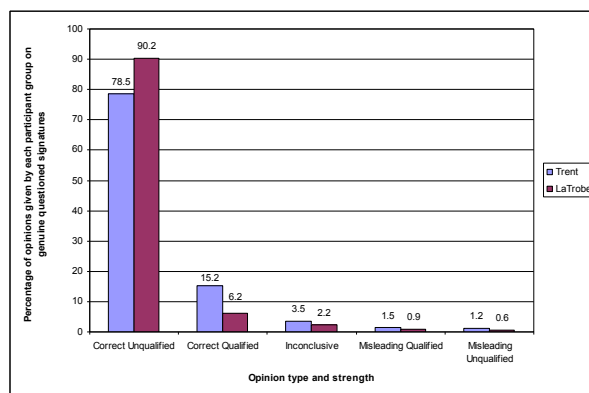


Figure 3 Comparison of Trent and La Trobe FDEs by opinion level for genuine questioned signatures.

($F = 40.59111$, $p = 0.0000$) given for simulated questioned signatures (see Table 9 and Figure 2).

Opinion Strength

When expressing an authorship opinion participants were instructed to choose their answer on a 5-level scale. Opinions at level 3 were 'inconclusive' and, as such, not considered to be either correct or misleading. Opinions 1 and 2 both indicated that the questioned signature was written by the specimen writer and were scored as a

correct opinion for genuine questioned signatures or as a misleading opinion for simulated questioned signatures. Opinions 4 and 5 indicated that the questioned signature was not written by the specimen writer and were scored as a correct opinion for simulated questioned signatures or as a misleading opinion for genuine questioned signatures. Any non-inconclusive opinion was therefore scored as either correct or misleading but, at the same time, each answer also reflected the participant's level of certainty or confidence in the call being made. Levels 1 and 5 are considered to be unqualified opinions which reflect a very high level of confidence while 2 and 4 are considered to be qualified in nature, reflecting a slightly less confident opinion.

A comparison and breakdown of answers, by opinion level and by type of signature, for both Trent and La Trobe FDEs is provided in Figures 3 and 4. Overall, there was no significant difference between the two FDE groups for any of the five levels or by type of signature (genuine vs. simulated). From the graph in Figure 3 it can be seen that there was a slight difference between the two groups in the number of 'correct unquali-

fied' and 'correct qualified' opinions for genuine questioned signatures. The FDEs in the Trent study appear to be slightly more conservative than those in the original La Trobe study; however, these results were not found to be statistically different.

Discussion

Trent Study versus 2007 La Trobe Study

The overall correct, misleading and inconclusive opinions given by the FDE participants in this study were compared to the results from the 2007 La Trobe signature comparison study (Found and Rogers 2007(2)) which included specimen and questioned signatures from the same signature set. FDEs in the Trent study achieved a correct opinion rate of 90.94% with an error rate of 1.75% and an inconclusive rate of 7.31%. FDEs in the La Trobe study gave 94.6% correct opinions with an error rate of 1.8% and an inconclusive rate of 3.6%. Although these values are slightly different, a statistical comparison of the results showed no significant difference between the two studies with respect to any of the rates.

By excluding the inconclusive opinions given by each participant we were left with their called opinions rates. Called opinions are based on the number of correct and/or misleading opinions made by participants when they were prepared to conclude, at least conditionally, that a signature was either genuine or simulated (Sita, Found, and Rogers 2002). In other words, participants provided a called opinion whenever they gave something other than an inconclusive opinion. No difference was found to exist between the called opinion rates of the Trent FDEs and the La Trobe FDEs. Trent FDEs had a rate of correct called opinions of 98.11% and a rate of misleading called opinions of 1.88%. The La Trobe FDEs had a rate of correct called opinions of 98.13% and a rate of misleading called opinions of 1.9%.

The similarity in the results between the two studies not only gives validation to the previous study since the earlier results were replicated, but also provides support for online testing as a viable, alternative testing method.

Forensic Document Examiners versus Non-Examiners

Based on previous studies involving signature comparison tests (Kam, Wetstein, and Conn 1994; Kam, Fielding, and Conn 1997; Kam, Gummadidala, Fielding, and Conn 2001; Sita, Found, and Rogers 2002) it was expected that FDE participants would give significantly fewer misleading opinions regarding the authorship of questioned signatures than non-examiners. Some previous studies (Kam, Fielding, and Conn 1997; Sita, Found, and Rogers 2002) also found that FDEs gave a greater number of inconclusive opinions and a similar number of correct opinions when compared to the non-examiner group. The results from the present study were in agreement with the previous results with respect to misleading opinions, with FDEs giving significantly fewer misleading opinions than the non-examiners. However, FDEs gave a similar number of inconclusive opinions and significantly more correct opinions. The FDE group gave 90.94% correct authorship opinions and 1.75% misleading opinions. Both of these rates were significantly different from the non-examiner group which gave 74.19% correct authorship opinions and 14.44% misleading opinions. There was no significant difference between the inconclusive opinions given by FDEs (7.31%) and non-examiners (11.38%).

One possible feature of the present study that could account for the differences in results from the previous studies is the type of opinion scale that was used. A study by Sita et al. (2002) found that FDEs gave a greater number of inconclusive opinions and a similar number of correct opinions as non-examiners. Participants in that study were instructed to conclude that a signature was either genuine or simulated or that they were unable to say (inconclusive). This is different from the five-point scale that was used in the current study which allowed for intermediate opinions of "it is or is not likely that the questioned signature was written by the specimen signature". A more restricted set of opinion options in the Sita et al. study could have led to the increased number of inconclusive opinions made by FDEs. With fewer opinion options an inconclusive opinion might be expressed whenever there was any amount of uncertainty about the authorship of the signature. In the current study, the participant had the option of expressing an intermediate opinion reflecting that uncertainty rather than having to revert to the inconclusive opinion. In other

words, the presence of the level 2 and 4 opinion options might have shifted the results away from inconclusive, if only slightly.

Although these results are not entirely consistent with other previous studies, the results support the argument that expertise for signature identification exists within the field of forensic document examination. In the present study, FDEs gave a greater number of correct opinions while making significantly fewer errors than the non-examiner participants.

The evidence in support of a claim of expertise was also seen when called opinions were considered. The FDE group had a significantly higher rate of correct called opinions and a significantly lower rate of misleading called opinions than the non-examiner group. This again supports the belief that FDEs have an expertise in signature comparison that non-examiners do not possess.

Genuine and Simulated Questioned Signatures

The questioned signature set was composed of genuine and simulated signatures. Previous studies have found differing results regarding the proficiency of FDEs with genuine versus simulated questioned signatures. Some studies have found that FDEs gave fewer misleading opinions for simulated signatures (Sita, Found, and Rogers 2002) while others found FDEs were better at giving authorship opinions for genuine signatures (Found and Rogers 2003; Found and Rogers 2005). In the present study FDEs performance was not affected by the type of questioned signature being considered. There was no significant difference in the number of correct, misleading, and inconclusive opinions given for genuine and simulated questioned signatures. This was also true for FDEs who participated in the 2007 La Trobe study. When the opinion rates for genuine and simulated signatures were compared for each group of FDEs, no significant difference was found (Figure 1).

A study by Found and Rogers (2005) reported that more misleading opinions were made with simulated questioned signatures than genuine questioned signatures. However, that particular study included disguised signatures as well as genuine and simulated signatures in the questioned signature set. The disguised signatures were written by the specimen writer in a way intended to hide their identity by altering the handwriting features. Including disguised signatures in their study added an extra element

to the signature comparison process, making the decision more realistic and more challenging. In the present study, as well as in the 2007 La Trobe study, if a participant did not believe the questioned signature was genuine, then it could be safely assumed that it was a simulation. When disguised signatures are introduced into the questioned set the participant must make a further assessment as to whether the non-genuine sample is a disguised signature by the specimen writer or a simulation by someone other than the specimen writer. As a result, this additional possibility makes the decision process more complex and potentially more difficult. An increased difficulty should logically cause participants to give more inconclusive opinions with questioned signatures.

In the present study, non-examiner opinions were affected by the type of questioned signature under consideration. That is, the non-examiner participants performed significantly better with genuine questioned signatures, giving more correct opinions and fewer misleading opinions than with the simulated questioned signatures. No difference was observed between the two types of questioned signatures in terms of inconclusive opinions.

Compared to the FDE group, the non-examiners gave fewer correct opinions and more misleading opinions for the simulated signatures. There was no significant difference between the two groups of participants with respect to the genuine signatures or the number of inconclusive opinions given for simulated signatures. These results indicate that the difference between the FDEs and non-examiners is primarily the result of a better performance by FDEs with respect to their calls on simulated questioned signatures.

Opinion Strength

The participants in the present study, as well as the 2007 La Trobe study, were asked to use a five-point opinion scale for their authorship opinions. Though five opinion options were present on the scale, only three results (correct, misleading or inconclusive) were possible. This is because opinion levels 1 and 2 both indicated that the participant believed the questioned signature was written by the specimen writer and opinion levels 4 and 5 both indicated that the participant did not believe the questioned signature was written by the specimen writer. Though both opinions were scored equivalently, they represented a different

level of certainty in the mind of the participant. Levels 1 and 5 are at the ends of the opinion scale and represent a very confident opinion, while levels 2 and 4 represent a slightly weaker opinion.

In practice, most forensic document examiners use a five-, seven-, or nine-point scale to express their authorship opinions. The American Society for Testing and Materials (ASTM 2008) describes the terminology for expressing opinions by FDEs. This standard describes a nine-point opinion scale on which opinions at the ends of the scale (levels 1 and 9) represent an extremely high level of confidence indicating that the examiner has no doubt about the authorship of the writing. The levels in between the strongest opinions and the midpoint (inconclusive) indicate that the examiner believes the writing either was written by or was not written by the specimen writer however some limitations (such as the quality or quantity of comparable writing) exist that are introducing a level of uncertainty into the examination. In other words, the scale provides information about the examiner's opinion regarding authorship and their confidence in that opinion.

When the opinions given by the Trent FDEs and the La Trobe FDEs at each of the five opinion levels were compared there was found to be no significant difference between the opinions expressed overall and for genuine and simulated signatures individually. This can be seen in Figures 3 and 4 in the Results section.

Writer 11

When the test samples for the original 2007 La Trobe study were prepared, twenty-four individuals were asked to create simulated signatures to be included in the questioned signature set. The majority of these writers made 20 simulations in one sitting with the exception of writer 11 who was asked to write 20 simulations each day for a period of ten days, giving a total of 200 signatures.

The summary of results from the 2007 La Trobe study reported that two questioned signatures written by writer 11 that were included in the questioned signature set were the only two signatures in that study to attract more than one misleading response. These two signatures were the 84th and 184th simulation made by writer 11 and therefore would have been written after several days of practice. A third signature was also included in the La Trobe trial and was the 24th simulation made by writer 11. That signature received at most one misleading response

(summary material provided by La Trobe indicated which signatures received more than one misleading opinion; therefore, all others would have received zero or one misleading response). The 24th signature would have been written on the second day of practice when the writer had only slightly more experience writing the simulation than the other writers. The authors felt that these results suggested that more practice may improve the appearance of a simulation thus causing examiners to be misled (Found and Rogers 2007(2)).

The present study also included signatures written by writer 11. The questioned signature set used in the present study was randomly selected and included the 8th and 11th simulations by this writer, both of which would have been written on the first day of practice when the writer had no more experience than the other simulators. As such, these signatures were similar to the third signature in the La Trobe study that was mentioned above. Both of the questioned signatures in the Trent study elicited no misleading opinions. Due to the fact that no signatures by writer 11 written when he or she had had more practice were included in the present study, it is not possible to comment further on the issue of whether or not a 'practised' simulation presents a greater challenge in authorship determination.

One signature included in the Trent study did stand out as attracting a very high number of misleading opinions compared to the rest of the questioned signatures. Signature 2, which was written by writer 19 on his or her 9th attempt, received ten misleading opinions by FDEs. This was much higher than the other questioned signatures which received between zero and three misleading responses, including a second signature written by writer 19 which received two misleading responses. It is not clear why the results for this particular signature were so different from the rest of the set, however; it is possible that it is due to the format of the test. If the participants had been able to group the signatures during their examination and review their responses before submitting them to the researcher the characteristics that would have identified this signature as a simulation may have been more easily observed.

Statistical Outlier

One member in the Trent FDE group (participant 143629) gave a total of 43 inconclusive opinions.

Forty-two of these opinions were given for simulated questioned signatures and one was given for a genuine questioned signature. This participant gave the greatest number of inconclusive opinions and was also the only participant to give this type of opinion for every one of the simulated questioned signatures. The high number of inconclusive opinions is very unusual when compared to the mean number of inconclusive opinions given by FDEs which was only 3.66. Since this number of inconclusive responses is unusual, it is tempting to call this participant an outlier. In statistical terms, an outlier may be defined as any isolated data point that lays more than three standard deviations away from the mean (Sincich 1986). The mean and standard deviation for inconclusive opinions by FDEs is 3.66 ± 7.94 . A score of 43 inconclusive opinions is almost five standard deviations away from the mean, clearly making this participant a statistical outlier. However, even though this participant is a statistical outlier, their responses still represent their opinions as a professional FDE and therefore should be taken into account. The reason for this behaviour is unclear but it is possible that the examiner adopted a personal or organizational policy of expressing only inconclusive opinions when dealing with non-original questioned signatures when those signatures display features suggesting they might be non-genuine. Such a policy has been implemented in some laboratories.

Previous Experience with La Trobe Signature Trials

The signatures that were used in the current study were from a set that had been previously used in a 2007 signature comparison trial by the Forensic Expertise Profiling Laboratory at La Trobe University. The signatures were randomly selected from the same 'master' set of signatures however each study used different questioned signatures. Because the signatures on both tests came from the same 'master' set it allowed the results from the two tests to be directly compared. At the same time, this created the possibility that an FDE in the Trent study may have also participated in the La Trobe study and therefore would have been previously exposed to the signature set. Furthermore, simply having some experience with the La Trobe testing procedure could result in a better performance. Only six of the 32 Trent FDEs indicated that they had previously participated in at least one La Trobe

proficiency test. This does not necessarily mean that they participated in the 2007 study. The limited number of FDEs with previous experience did not allow for meaningful statistical testing, however the FDEs with previous experience did not appear to perform differently on the online test relative to examiners with no experience. The one exception to this was participant 143629 who was an FDE with previous experience with La Trobe tests. This subject's results did differ from the FDEs who had not participated in La Trobe testing but they also differed from those who had experience. It appears that the difference in results for this participant was due to some factor other than previous testing experience. More information on this participant is provided in the Statistical Outlier section.

Trainees

Two participants identified themselves as FDE trainees in their personal information. These individuals were included in the FDE group even though they are not yet professional examiners. Since they are training to be professional document examiners they would have had some experience with signature identification that members of the non-examiner group would not have had. Unfortunately, there was no way to determine where these individuals might have been in their training program and, therefore, no way to accurately measure if training was an issue in the testing procedure. However, no obvious effect was observed in the opinions of the 'trainees' which were similar to the opinions of the rest of the FDE group. For these reasons the two trainees were included as members of the FDE group.

The placement of participant 702009 into a particular group was also difficult. This participant indicated that they had very little experience with document examination yet they are associated with a forensic document examination group. All non-examiner participants were Trent University students, and therefore 702009 could not simply be moved into the non-examiner group. Since this participant was invited to participate as an FDE and they are associated with an FDE organization, they were placed in the FDE participant group.

Participation

Approximately 230 FDEs and 200 Trent University students were invited to participate

in the online signature comparison test. Out of those contacted, 32 from each group completed the test, giving a total of 64 participants. This represents a final response rate of only 13.9% for FDEs and 16% for non-examiners. There were also participants who submitted their registration information but did not begin the test itself, or began the test but failed to complete it. In total, 95 individuals submitted a registration but only 64 completed the signature comparison giving a drop-out rate of 32.6%. For both groups these are not high response rates and may limit the ability to generalize the results to the broader FDE and non-examiner communities.

There are several reasons that can be suggested for the small number of invitees who agreed to participate in the study. There is a chance that not all of the invitations successfully reached the FDEs and they therefore could not participate. The contact information of the FDEs who were invited to participate was obtained through the American Board of Forensic Document Examiners (ABFDE) and American Society of Questioned Document Examiners (ASQDE) websites, as well as the contact list from the 2007 European Network of Forensic Handwriting Experts (ENFHEX) meeting in Vilnius, Lithuania. If this contact information was incorrect or was not up to date the examiner would not have received the invitation and would not have been able to participate. Some individuals may have also been restricted from participating due to the fact that the invitations were sent via the researchers university email account. Email accounts often contain filters that do not accept emails from unrecognized senders. In the current study, the invitations were sent from a Trent University email account and may not have been recognized as an acceptable email address by many of the invitees, especially those associated with government agencies or from other countries. This would not be a problem with the non-examiner invitees since they were all students at Trent University and would have no problem receiving emails from another Trent email account.

Along with invitees who may not have been able to participate due to not receiving an invitation, examiners may have not participated because the test did not include original signatures. Since the test was presented online, the signatures were scans of the originals and were therefore not original themselves. A few examiners contacted the researcher after being invited

to participate indicating that the agencies they worked for did not support giving authorship opinions for questioned signatures that were not original. There may have been other invitees who did not contact the researcher but also did not participate for a similar reason.

Once individuals had received the invitation and decided that they would participate there are other factors associated with the design of the website that may have restricted their ability to complete the test. In order to properly view the signature images and complete the trial the participants were required to have Java software installed on their computer. Any invitee who did not have Java on his or her computer was unable to participate unless they were willing to install the software. Similarly, the study required the use of cookies which some participants may not have been willing or able to accept. For example, subjects trying to do the study from a work location might have been restricted due to security limitations on their internet connection or software.

Problems were also associated with the personal information page that each participant was required to complete before beginning the online test. One of the main problems associated with this page was that when a participant finished filling out their personal information and clicked on the submit button the website would often direct them back to the introductory page at the very beginning of the test instead of taking them to the first questioned signature page. The participant would then have to fill out their personal information again and resubmit it. This did not occur with every participant but in several cases the participant had to submit their information multiple times before being able to see the first question. Once this problem was identified a note was added to the personal information page advising participants that they may be required to enter their information more than once. If participants did not see the warning about the problem they may have assumed the website was not working properly and not continued on with the test. A second problem was encountered with the personal information page, this time involving the identification of the participant's certifying body. When participants registered their personal information they were asked if they were certified as a professional document examiner and if so by which certifying body. In some cases their selection of certifying body would not be accepted and the participant would

not be able to go on to the first testing question. The only way to proceed with the test was to select the ABFDE certification option. Once this problem was identified it was immediately corrected however a number of participants would have encountered it. Before the problem was corrected, some participants included a note in the comments section of the personal information page informing the researcher of their correct certifying body. However, there may have been some FDEs and non-examiners who encountered this problem and did not continue on with the examination.

Each test page included one questioned signature and 14 specimen signatures. All 15 of these images were fairly large JPEG files. The size and quality of these images caused them to load very slowly if the participant did not have a high-speed internet connection. The length of time it took to load the images may have caused some participants to discontinue the test. Though the researchers were aware of this possibility, it was more important to have high quality rather than fast loading images.

It is important to note that the functionality of the study had been tested extensively beforehand. Despite the pre-trial testing, several small but potentially significant issues were missed. This is a difficulty that is encountered with most web-based systems since it is very difficult to avoid or eliminate issues relating to user setup, software or connection restrictions.

All of these problems could account for the large number of individuals contacted who chose not to participate. All of these issues should be addressed in future studies that use an online presentation format.

Methodological Limitations

After completing all of the test questions, participants were given the chance to submit any comments they had about the online test. From these comments a list of limitations of the test that are associated with its layout were compiled.

The main criticism that was identified by several FDEs was the inability to compare the questioned signatures to each other before submitting their answers. The test was designed so that participants were required to submit their authorship opinion for each signature before continuing on to the next question. When a signature test is performed using the traditional paper method, the participants are able to view

all of the questioned signatures at once. This allows for the questioned signatures to be compared to each other and to be grouped based on their similarities.

The suggestion was made that since participants were not able to view all of the questioned signatures before they were required to give an opinion they should be allowed to review their opinions at the end of the test before submitting them to the researcher. Therefore, if over the course of the test the participant decides that they incorrectly identified a signature as being genuine or simulated, they would have the opportunity to change their opinion. Requiring an opinion to be made about each questioned signature before moving on to the next questioned signature does not invalidate the signature comparison test. It does, however, change the format from a comparison of a set of questioned signatures to a test that consists of several individual comparisons performed sequentially.

Though these factors were identified as limitations by multiple FDE participants they did not seem to affect their overall ability to give authorship opinions. FDEs who participated in this study gave a similar number of correct, misleading and inconclusive opinions to FDEs in the previous La Trobe study which was a more traditional paper version of the same test. It was suggested in a previous section that some elements of the test's design may have contributed to the large number of misleading responses made on Signature 2. This, however, cannot be explored without further testing using a format that allows for grouping and review of questioned signatures.

Another limitation was the inability to view the instruction page once the participant had begun the test. Several invitees, including all of the FDEs, were sent the test instructions as part of their invitation. Unfortunately, individuals who were not invited to participate by email, such as students who were invited through class announcements, were not given the instructions ahead of time. The ability to review the instructions throughout the test could have been especially important for non-examiner participants and those who had never completed a signature comparison test. It may have been beneficial for these participants to be able to review the instructions once they had started the test. It is possible that some participants would have read the instructions quickly, not realising that they would not be able to see them again. If the participants did not completely understand the

task, their answers may not represent their true opinions.

Some participants suggested that they would have liked the option to print out the signatures for comparison. This is an option that could be considered in future studies. However it should be noted that this introduces another potential variable to consider because it becomes impossible to account for potential differences in the resulting print quality which might affect the decision making process.

Positive comments about the online test were also received from participants. A few participants commented on the high resolution of the questioned and specimen signatures and felt they were of similar quality to those included in previous tests. Other participants mentioned that the website was easy to use and the test itself was easy to understand.

Future Recommendations

The limitations outlined by participants in this study should be taken into consideration in future online proficiency tests. In order to make the online test more similar to paper-based tests as well as casework the participants could have access to all of the questioned signatures before having to make a decision as to whether they were written by the specimen writer. The comments from the FDE participants suggest that they would appreciate being able to view all of the questioned signatures at the beginning of the test, though the results show that this did not affect their ability to give an opinion. Along with being able to see all of the questioned signatures before beginning the test, participants could also be allowed to review their opinions before submitting them to the researcher. Since this format would better reflect casework situations, there is also the possibility that a greater number of FDEs would be willing to participate.

In the present study the participant's test responses were sent to the researchers email account immediately after being submitted by the participant. While this allowed for the responses to be received quickly, it did not make the data input or analysis any easier nor did it eliminate the possibility of transcription error during data entry. Future studies could design the website so that data is automatically entered in a database, spreadsheet or statistical program to eliminate the process of inputting the data and to facilitate data analysis.

It should be possible to make use of server functionality to better control the study without relying upon user settings or client-side software. For example, a database-driven system requiring user registration and login could be used to provide all of the above-mentioned functionality as well as collection of better biographic data without reliance on java or cookies.

A signature comparison test was completed in this study, but this is just one of the areas of forensic document examination that could be tested. Further studies should be completed looking at the use of an online test for other types of comparisons such as handwriting and handprinting, as well as signature tests involving disguised signatures. The fact that FDEs performed similarly on the two types of tests for signature comparison does not guarantee that they will also perform similarly when comparing other types of writing. However, since the online signature test was successful it suggests that an online testing format would be possible for other types of writing.

Conclusion

The present study was designed to test a new online method for performing proficiency testing by comparing the results of FDEs to those collected from a more traditional paper form of the test. A statistical comparison showed that there was no significant difference between the authorship opinions given by the FDE participants in present study and those from the 2007 study by the Forensic Expertise Profiling Laboratory at La Trobe University. Both FDE groups gave a similar percentage of correct, misleading, and inconclusive opinions, as well as correct called and misleading called opinions. Similar opinions were also given by FDE participants in the Trent and La Trobe studies for genuine and simulated signatures as well as for the different levels of authorship opinions. These results support the use of the online method as an alternative form of proficiency testing in the field of forensic document analysis.

A group of non-examiner participants were included in this study to determine if the results from the online test would support previous findings that the abilities of professional FDEs are superior to those of the average person. The results of the two groups of participants revealed that FDEs performed better than non-examiners by providing a greater number of correct author-

ship opinions and fewer misleading opinions. It was also found that the opinions given by FDEs and non-examiners were significantly different for simulated signatures but were not different for genuine signatures. These results indicate that the difference between the FDEs and non-examiners is primarily due to the better performance of FDEs with simulated questioned signatures. The significant difference between the opinions of the FDEs and non-examiners gives further support to the finding that an expertise exists in the field of forensic signature identification.

References

- American Society for Testing and Materials International (2008). Standard Terminology for Expressing Conclusions of Forensic Document Examiners, ASTM E1658-08.
- Found, B. and Rogers, D. (Eds) (2003). Documentation of Forensic Handwriting Comparison and Identification Method: A Modular Approach. Revised version. La Trobe University, Australia. Originally published in the *Journal of Forensic Science* (1999), Vol. 12, pp. 1-68.
- Found, B. and Rogers, D. (2005). Problem Types of Questioned Handwritten Text for Forensic Document Examiners. Proceedings of the 12th Biennial Conference of the International Graphonomics Society, pp. 8-12.
- Found, B. and Rogers, D. (2006) (1). Revision and Education Package: Handwriting Proficiency Trial 2006. Forensic Expertise Profiling Laboratory, School of Human Biosciences, La Trobe University, Australia.
- Found, B. and Rogers, D. (2006) (2). Production notes and samples for 2007 Signature Trial, Forensic Expertise Profiling Laboratory, School of Human Biosciences, La Trobe University, Australia.
- Found, B. and Rogers, D. (2006) (3). Revision and Education Package: Signature Trial 2006. Forensic Expertise Profiling Laboratory, School of Human Biosciences, La Trobe University, Australia.
- Found, B. and Rogers, D. (2007) (1). Revision and Education Package: Handwriting Proficiency Trial 2007. Forensic Expertise Profiling Laboratory, School of Human Biosciences, La Trobe University, Australia.
- Found, B. and Rogers, D. (2007) (2). Revision and Education Package: Signature Trial 2007. Forensic Expertise Profiling Laboratory, School of Human Biosciences, La Trobe University, Australia.
- Found, B. and Rogers, D. (2008) (1). Revision and Education Package: Handwriting Proficiency Trial 2008. Forensic Expertise Profiling Laboratory, School of Human Biosciences, La Trobe University, Australia.
- Found, B. and Rogers, D. (2008) (2). Revision and Education Package: Signature Trial 2008. Forensic Expertise Profiling Laboratory, School of Human Biosciences, La Trobe University, Australia.
- Galbraith, O.; Galbraith, C.; and Galbraith, N. (1995). The Principle of the "Drunkard's Search" as a Proxy for Scientific Analysis: The Misuse of Handwriting Test Data in a Law Journal Article. *International Journal of Forensic Document Examiners*, Vol. 1, No. 1, pp. 7-17.
- Hammond, D. and Found, B. (2008). Post-Test Information Package: Skill-Task Assessment #S-01 (2008). Skill-Task Training Assessment and Research, Inc., United States and Forensic Expertise Profiling Laboratory, School of Human Biosciences, La Trobe University, Australia.
- Hammond, D. and Found, B. (2009). Post-Test Information Package: Skill-Task Assessment #S-02 (2009). Skill-Task Training Assessment and Research, Inc., United States and Forensic Expertise Profiling Laboratory, School of Human Biosciences, La Trobe University, Australia.
- Kam, M.; Fielding, G.; and Conn, R. (1997). Writer Identification by Professional Document Examiners. *Journal of Forensic Sciences*, Vol. 42, No. 5, pp. 778-786.
- Kam, M.; Gummadidala, K.; Fielding, G.; and Conn, R. (2001). Signature Authentication by Forensic Document Examiners. *Journal of Forensic Sciences*, Vol. 46, No. 4, 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, No. 1, pp. 5-14.
- Kelly, J.S. and Lindblom, B. (Editors) (2006). *Scientific Examination of Questioned Documents*, 2nd ed. Boca Raton, FL: CRC Press, pp. 10-12.
- Risinger, D.M.; Denbeaux, M.P.; and Saks, M.J. (1989). Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lesson of Handwriting Identification "Expertise". *University of Pennsylvania Law Review*, Vol. 137, pp. 731-792.

Sincich, T. (1986). *Business Statistics by Example*. 2nd ed. San Francisco: Dellen.

Sita, J.; Found, B.; and Rogers, D. (2002). Forensic Handwriting Examiners' Expertise for Signature Comparison. *Journal of Forensic Sciences*, Vol. 47, No. 5, pp. 1-8.