Journal of the American Society of Questioned Document Examiners, Inc.
Volume 17, Number 2, December 2014

Investigating Forensic Document Examiner Opinions on Signatures in a Foreign Script[1,2]

Rosalind Spencer

Canada Border Services Agency, Ottawa, Canada. Email: Rosalind@whatiftechnologies.com

There have been court challenges in various countries around the world bringing into question the validity of forensic document examiners (FDEs) expressing expert opinions on writing systems other than those with which they are familiar. A study was designed to assess whether FDEs are significantly better at evaluating the authenticity of questioned signatures in a familiar script than FDEs who are faced with an unfamiliar one. Volunteer FDEs familiar with Chinese script, those unfamiliar with the script, and a group of laypeople were provided with a series of small validation tests in a blinded fashion. The theoretical paradigms used by FDEs to form the basis of expert opinion in this area were explored, together with a discussion on how these results may be generalised to the overall FDE population, given the limitations of such a study.

Introduction

In 2009 the NAS report was published entitled, "Strengthening Forensic Science in the United States: A Path Forward." Its aim was to bolster the science behind the forensic sciences as practiced in the United States. It also noted that the development and application of the forensic science disciplines to support intelligence, investigations, and operations aimed at the prevention, interdiction, disruption, attribution, and prosecution of terrorism has become an important component of both public health and what is now termed "homeland security" over the past two decades. With the development and deployment of enhanced capabilities has come the integration of forensic science disciplines much earlier in the investigative process. As a result, the forensic science disciplines are being leveraged to generate investigative leads to test, direct, or redirect lines of investigation, as distinct from building a case for prosecution.

The report included 'Questioned Document Examination', i.e. the examination of signatures, handwriting and hand printing as practiced by formally-trained forensic document examiners (FDEs) such as those of the American Society of Questioned Document Examiners (ASQDE), as an area of concern because these examinations often lead to identification of a specific individual. It is acknowledged that many of the existing forensic document examination (FDE) technical papers and even some textbooks will fall short of the scientific template being applied retroactively to their content. Attorneys and the courts will rightly demand validation studies and empirical research in our field.

The trends of globalisation within social, economic and forensic communities, and greater scrutiny of forensic evidence in court, are currently combined with a relatively small number of properly trained FDEs around the world, who may be asked to opine on many different types of signatures. Blending of national identities is on the increase, either by the creation of large immigrant populations in countries such as Canada and Australia, or by mass migration as is occurring currently in Europe. This means that the traditional 'class vs. individual characteristics' and 'complexity' paradigms are coming into play in a different way. In theory, it follows that if complexity provides an adequate platform on which to assess authorship, then 'familiarity' i.e. language fluency, should not be of concern. The extent to which errors may or may not be compounded when languages foreign to the FDE are examined has not been extensively studied.

[1]All comments in this paper are those of the author and do not necessarily represent those of the Canada Border Services Agency and/or the Government of Canada.
[2]The review process for this paper was conducted by the Assistant Editor.

There is a need for a cohesive response by the profession to this type of challenge.

One area in which there has been ongoing debate is whether it is necessary for a forensic handwriting examiner to have a complete knowledge of the language of the written material in question. This query cropped up again in 2012 in an international forum for forensic document examiners[3], and generated an interesting and varied set of responses from a range of individuals who either did or did not render opinions on handwriting or signatures in foreign scripts. Views expressed depended on whether handwriting or signatures were in question. Requests were made for any literature to convince the courts regarding acceptance of such an expert opinion. It became apparent that there was a need for more research in this area, covering a variety of different non-Latin scripts. Since the examination of signatures, as opposed to handwriting, in foreign scripts generated the hottest debate, it is this area on which the present study focuses.

One of the earliest references on this subject is Brewester, who in 1932, when referring to examiner knowledge of the questioned writer's language, opined that 'no such knowledge is necessary when dealing with signatures'. At the same time, he maintained that the examiner should have made an intensive study of the manner in which the characters are made, in order to follow correctly the sequence of each outline from beginning to end. Without such study Brewester deemed it impossible for an examiner to discover if a character were made the wrong way round, relative to the norm. He ended by saying 'This is an occurrence that often happens in a fraudulent signature.' Given the characteristic patterns of stroke order in Chinese ideographic writing (Johnston, 2012), which contrast with the multitudinous ways of constructing many Latin characters, this statement may be less applicable. The stroke order would in all likelihood be very familiar to a person fluent[4] in Chinese executing a simulation, and deviation from the norm might instantly raise suspicion.

Crown & Shimaoka (1974), and later Mittal & Sehgal (1989) recommended the seeking out of a linguist with knowledge of a particular script in the first instance, and postulated that an expert with knowledge of an allied (closely related) script could attempt a comparison, since there is reasonable familiarity and the problem of comparison is limited in nature. However these recommendations appear more relevant to examination of handwriting rather than signatures.

Published studies involving statistical analysis of the ability of FDEs to determine genuine signatures from signature simulations in non-Latin scripts are few and far between. The success and validity of such studies usually depends upon voluntary participation by FDEs, particularly those fluent in the script in question, and finding such participants is often a challenge and a limitation. A recent study investigating this question using Arabic signatures (Al-Musa Alkahtani, 2010), involved four FDEs, two fluent in Arabic and two unfamiliar with the Arabic language. Arabic is quite different from Roman script – whilst both are linear, cursive and derived from the Middle East, the letters are completely different. It is not easy to guess where the boundaries between letters may be and the writing is in different directions across the page. Slant and alignment also tend to be different. Al-Musa Alkahtani found no significant differences between fluent and non-fluent FDEs in their ability to assess line quality, form or proportion in a simulation.

A need was identified for a statistical study involving a non-Latin script, if possible on a larger scale than that of Al-Musa Alkahtani. Chinese ideographic writing was chosen for the following reasons. Firstly there has been a rapid increase in travel, dialogue and business dealings between East and West in recent years. Secondly cases involving this script are regularly faced by Western FDEs. Thirdly, ideographic Chinese script is completely different from both Arabic and Roman script. This study set out from the start to involve as many FDEs fluent and non-fluent in Chinese as practicably possible given certain constraints. These included cost, time and availability of FDEs.

Forensic examination of ideographic writing involves the same techniques used in the analysis of Latin script, and as with alphabetic writing, an individual can write with less skill than he possesses, but cannot write with a higher order of skill (Crown & Shimaoka, 1974). Similarly, in ideographic writing, a number of aspects may be

[3]Docexam_L email list maintained by Chris Anderson & Co Pty Ltd, Australia

[4]In the present study, fluent was defined as 'able to easily identify, read, write and sign in a particular language'. It may also be thought of as 'familiarity' with a language.

indicative of intentional disguise, such as a mixture of well and poorly-written characters, inconsistency in speed and style of writing, retouching, and wide variance.

To a Western eye, Chinese characters bear an overall simplicity yet bewildering multiplicity in detail. The most comprehensive dictionaries list almost 50,000 characters, comprising current and obsolete signs (Coulmas, 1999). These Chinese characters were not formed arbitrarily. Over several thousands of years, early religious writers, court officials, scholars, poets and eventually lexicographers created a body of material that shows a remarkable internal system. A few characters were originally pictographic; a few were arbitrary; but more were the result of combining phonetic elements with semantic features (Wilder & Ingram, 1974). During this time there have been many transformations for ease in writing or condensations for the sake of speed. Many of these changes occurred as writing materials and instruments changed. The Chinese writing system combines the representation of sound and meaning, with efficiency and speed, in a visual manifestation of language (Coulmas, 1999). This literacy made China the dominant culture in East Asia, and Chinese characters were borrowed to write other languages, notably Vietnamese, Korean and Japanese.

In modern times the Chinese script has undergone some modification as a result of the character simplification carried out in the People's Republic of China. Many abbreviated varieties of characters have become the standard forms (Coulmas, 1999). The reform programme of the 1950s and 1960s defined three main tasks: namely the standardization and simplification of Chinese characters; the creation of a Romanized orthography; and the promotion of the standard language. Simplified Chinese is standard in the People's Republic of China, Singapore, and Malaysia. Traditional Chinese is retained in Hong Kong, Taiwan, Macau and some overseas Chinese communities (except Singapore and Malaysia). Further simplification may not occur in the foreseeable future. The second simplification happened in 1977 after the first simplification in 1956, but ceased within a year because it was viewed as a threat to traditional Chinese culture, which is cherished much more today than in the past.

In Chinese communities handwriting is taught in primary school. Each character is a word, unlike Roman script. There are some broad differences in the way in which writing is taught in schools in the West and in China. Firstly, Chinese calligraphy is revered in the Chinese cultural sphere, and although less widely practiced in recent years due to the use of computers, most young students have some exposure to it. These enhanced writing skills may be retained throughout life, and some individuals continue on to perfect their calligraphy as an art form. Secondly, left-handedness in writing is still generally discouraged in China from an early age, and children are made to use their right hand.

Literacy rates in China are roughly 92.2% overall although there are differences between regions and the sexes. For example, 96% of males and 88.5% of females of age 15 and over can read and write[5], compared with 99% of either sex in the US and UK. Literacy in China and other countries where Chinese is an official language is graded on the number of characters in the speaker's lexicon, with a few thousand considered the minimum for practical literacy.

Generally speaking, in Chinese writing, individual Chinese characters are typically written with a particular stroke order, which may have the effect of making the character look more balanced (Johnston, 2011), but also aids a higher writing speed. About 9,000 words have standardized stroke orders (these words being most frequently used in handwriting, and thus of less value for handwriting identification). Every child in primary school learns the standardized stroke order when learning a word. Incorrect stroke orders are not allowed when a child learns to write in primary school. However, to a greater or lesser extent, for every individual some words are inevitably developed with non-standard stroke orders. These represent valuable features for FDEs since they are unique to an individual, especially when they occur at lower frequencies.

There is no standard stroke order for cursive writing. The stroke orders of handprinting and cursive writing of the same word are often different in one person's writing. So it is common that a handprinted signature differs dramatically from an individual's cursive signature in terms of stroke order. It may all depend upon the individ-

[5]United Nations Educational, Scientific and Cultural Organization (UNESCO) and UNESCO/UIS (UNESCO Institute of Statistics), including the Education for All 2000 Assessment. 2011 data.

ual and the degree to which a 'freeform' signature style is used.

In general, in a Chinese signature, the family name is written first (and pronounced first), followed by the individual's name, in a horizontal alignment from left to right, or in a vertical alignment. Writers may exhibit connections between these different characters to a greater or lesser degree. Cursive signatures are very common, and in these, two or three words are always connected together. The proportions and size of signatures may vary depending on the degree of constraint, in common with Western-style signatures (Simsons *et al*., 2011, Fazio, 2011, Desjardins, 2012). Writing instruments in present day China are commonly gel pens and roller pens (ballpoint pens are less popular), with brushes used for calligraphy. Paper is similar to that in the West. Fountain pen nibs more fully express stroke width variation, and may be used by businessmen solely to enhance the appearance of their signatures. The same features which form part of a forensic examination in Latin script signatures, such as variable pen pressure, speed and fluency, can be applied to an examination of Chinese script signatures. Examples of these are illustrated in Figure 1.

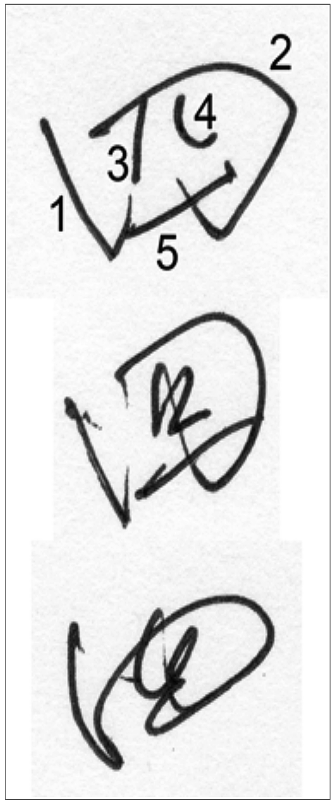

Figure 1. Examples of variable pen pressure, speed, fluency and similar stroke order (denoted by numbers) in different versions written by one writer, of the handwritten Chinese numeral character '4'.

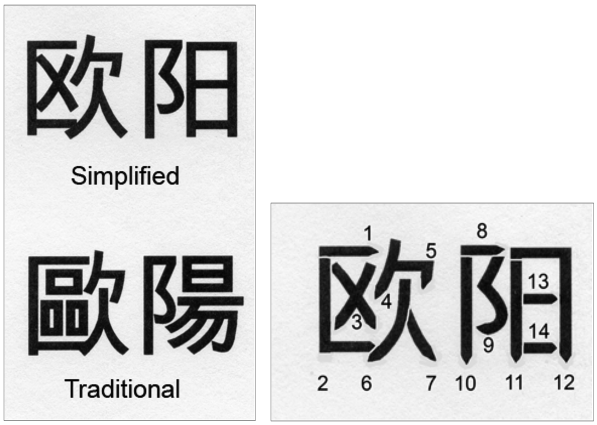

Figure 2. Example of the family name 'Ouyang' in simplified vs. traditional Chinese script, with usual order of strokes of the simplified version (numbers).

Figure 2 is an example of the family name 'Ouyang' in both simplified and traditional Chinese script. The usual order of strokes of the simplified version is also illustrated. In addition to the geographical areas which use traditional Chinese for all writing, some individuals who use simplified Chinese in everyday writing may use the traditional style for signatures, or vice-versa. Both styles may be used in the same sentence. It is not uncommon to alternate between the two styles for signatures. The traditional style is deemed more beautiful, and preferred by those who love calligraphy.

Methods and Materials

Participants

The experiment was designed from the outset with input from the Canada Border Services Agency (CBSA) Math & Data Exploration Section to allow for reliable statistical analysis of results, with an internal design for it to be randomisable where possible, and repeatable. Participants were drawn from a list of international volunteers, solicited by means of announcements to members of the ASQDE, American Board of Forensic Document Examiners, American Academy of Forensic Sciences, Canadian Society of Forensic Science, Australian and New Zealand

Forensic Science Society, CBSA, Docexam_L, local University International Student groups and local communities. Whilst these represent stratified samples of largely above-average education, in mostly technical fields, they were deemed the most efficient methods of reaching a large pool of potential volunteers. Study material was sent to participants in both high quality photographic and electronic form.

Participant questionnaire

This study included naturally written, simulated and disguised signatures. FDEs and laypeople familiar with Chinese script, and those unfamiliar with that script, were provided with a series of small validation tests in a blinded fashion. Background demographic and other information such as fluency in the Chinese language was ascertained by means of a brief participant questionnaire (Appendix 1), and any information provided was strictly anonymised. Participants were asked whether they had studied Chinese calligraphy, since in China calligraphy and enhancement of writing skills is encouraged, and skilled calligraphers have been found to simulate handwriting well enough to fool experienced FDEs (Dewhurst *et al.*, 2008). The preponderance of left- or right-handedness in writing was compared with previous studies (Leung[1] *et al.,* 1993).

The theoretical paradigms ('class-individual' versus 'complexity') used by FDEs to form the basis of expert opinion in this area were explored.

Experimental Protocol

Collection of test material

With every participant, the project was introduced, with an opportunity for questions, and then two copies of a signed consent form were obtained, one retained by the participant. These consent signatures were treated as additional known signatures but were not used as masters for simulation purposes. A participant questionnaire was also completed at this time, and a unique signer number assigned to the individual. From a group of fluent Chinese-speaking student volunteers, 10 individuals were randomly picked as a sample set. From these students, 132 naturally-written known specimen signatures (12 per individual, except an ambidextrous individual who gave 24) were collected in one sitting on specially designed forms (Appendix 2). The signature collection and random sampling procedure is summarised in Figure 3. The paper used was Domtar Earth Choice office paper, Item 2701, White 92, Letter size, Weight 20lb, 500 sheets, 3-hole punch, with Paper Mate Flexigrip Ultra blue ballpoint 1.0mm (M) pens. Five of the genuine signature names were randomly chosen to be used in the pool of test K samples.

The 10 individuals chosen by simple random sampling without replacement, were picked to provide 12 simulations of these test known (K) signature names in a freehand way, after practicing on a plain sheet of letter-size paper with a close-up view of the genuine ink signatures for as long as they felt necessary (about 5-10 minutes in most cases). This produced 120 simulations of each name. A mixed subset of 10 genuine and simulated signatures, was randomly chosen as the test questioned (Q) samples.

In a random order for each name, both Q and K samples were scanned at 1200ppi, stored on discs and printed on quality photographic paper at actual size. The test material was distributed to volunteer participants for evaluation, together with full instructions (Appendix 3), the participant questionnaire and answer sheet, with generous time allowed for completion. The answer sheet, shown in Appendix 4, restricted participants to the options of giving either a definite or a less-than-definite opinion as to whether a known writer wrote or did not write a particular Q signature (forced calls). No formal opinion scale was used, simply to make it applicable to a wide range of FDEs and laypeople. No inconclusive opinions and no peer review were allowed. Results were received by hand, email or post. Personal information was stripped from the questionnaires upon receipt, and images were not published elsewhere for privacy reasons.

Participants

The participants fell into five different blocks, as follows:

- Fluent in Chinese, FDE
- Non-fluent in Chinese, FDE, report on Chinese signatures
- Non-fluent in Chinese FDEs, don't report on Chinese signatures
- Fluent in Chinese, non-FDE
- Non-fluent in Chinese, non-FDE

The broad geographical regions from which

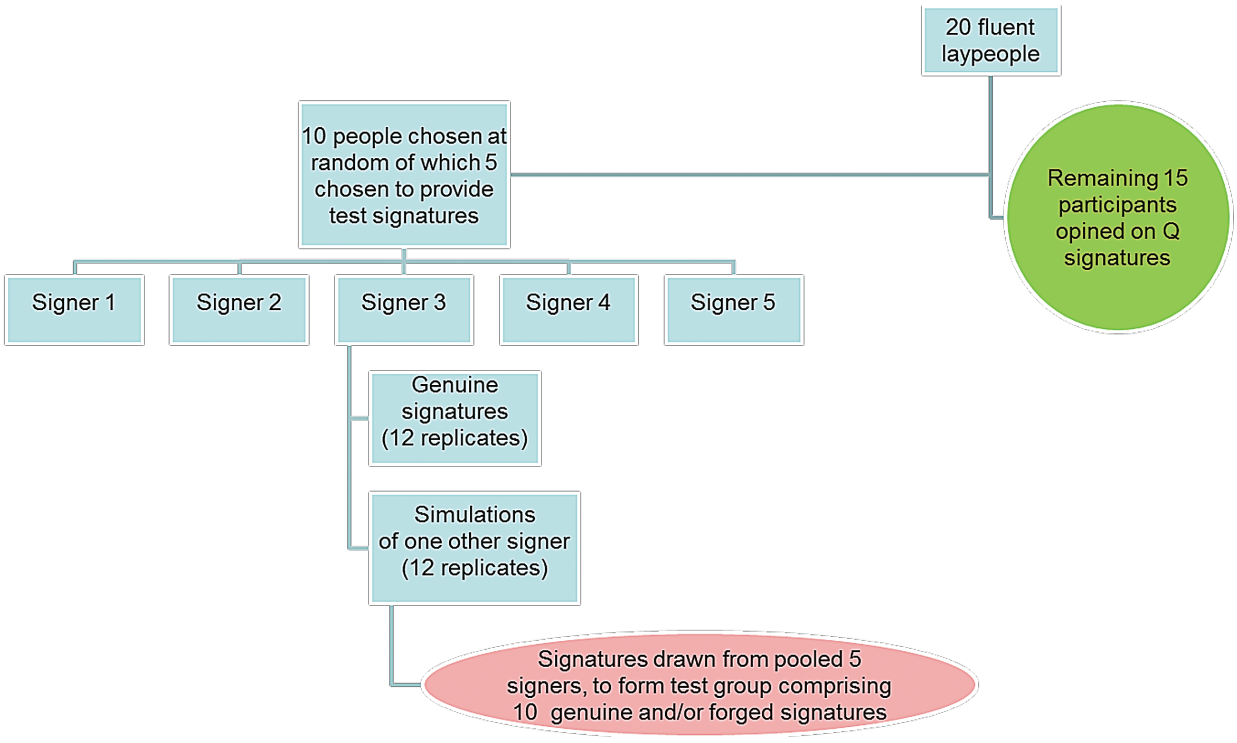

Figure 3. Overview of test data collection. Each of signers 1 to 4 behaved similarly. Ambidextrous signer 5 provided 24 genuine signatures and 12 simulations.

participants were drawn are shown in Appendix 5. Laypeople ('non-FDE') were chosen from areas to mirror as much as possible the international context of the FDEs. Fluent laypeople and fluent FDEs came mostly from the Orient, whilst non-fluent laypeople and non-fluent FDEs came from countries in the West. Fluency was self-identified in the participant questionnaire.

Hypotheses

A number of hypotheses were tested to try to answer the following questions:

1. Is there a difference between the ability of an FDE who is not fluent in Chinese and reports on signatures in Chinese script, and the ability of an FDE not fluent in Chinese who does not report on signatures in Chinese script, to assess authorship of a signature in Chinese script?
2. Is there a difference between the ability of an FDE who is fluent in Chinese, and the ability of an FDE not fluent in Chinese, to assess authorship of a signature in Chinese script?
3. Is there a difference between the ability of an FDE who is fluent in Chinese, and the ability of a layperson fluent in Chinese, to assess authorship of a signature in Chinese script?
4. Is there a difference between the ability of an FDE who is not fluent in Chinese, and the ability of a layperson not fluent in Chinese, to assess authorship of a signature in Chinese script?
5. Are there differences between the above assessment abilities depending on which questioned signature is assessed?

The study sample comprised all subjects willing and able to participate in the study, and the theoretical target population was all formally-trained FDEs in the world. Factors considered in the experimental design included the randomisation of as many variables as possible, to make the sampling representative of the general population of interest, given logistical constraints. Elimination of some obvious sources of bias was achieved by all participants receiving the same samples, al-

though differences in motivation may have been a factor not accounted for. Not all laypeople had technical or scientific backgrounds, and the spread of ages was fairly broad. The five blocks are mutually exclusive groups. However, the five groups are not random groups, since the very fact of volunteering means that they are stratified. The study sample is a convenience sample and self-selected.

The study was designed to be as close to real life as possible, whereby an FDE does not know all the circumstances of how the signatures came to be written, whether there was disguise or simulation involved. If information was requested, participants were directed that any circumstance was possible. Disguise was not actively excluded from the non-simulated Q samples, but coincidentally a form of disguise came to be included. The signatures were chosen randomly from the donors, and fortunately included a range of complexity; traditional ideographic characters vs. stylised forms; and language legibility or recognisability, which was desirable to mimic real life. Serendipitously one of the donors was ambidextrous, and provided specimens in both hands. When samples by opposite hands were randomly chosen as Q and K, it was decided to keep them in the trial, since disguise had not been excluded, even though there was no intention on the part of the writer to actively disguise their signature. This individual used the right hand for writing and the left hand for other daily tasks. No peer review was allowed, even though in their normal course of work, many FDEs would not produce a report without doing so.

Statistical analysis

It was decided to analyse the results in two different ways. Firstly, for each participant, one out of their 10 answers was randomly chosen to be used in the analysis (Analysis 1). The aim was to average out bias between people and signatures, but inadvertently may have had a deleterious effect by overly diluting the data. Secondly, each pair of Q signatures in the same name was analysed separately, across all groups. This would account for any 'problematic' Q signatures.

Results

The response rates to the study on receipt, were 99.9% and 95% for fluent FDEs and non-fluent FDEs respectively. In terms of numbers, the response from FDEs fluent in Chinese was outstanding, and that of 'Western' FDEs was considerably lower. However, extrapolating from the numbers of FDEs in Jiangsu Province, it is believed there may be approximately 2,700 formally-trained forensic handwriting examiners in China today, and around 1000 in the 'Western' world. If so, then the participation rates of fluent and non-fluent FDEs in this study were fairly similar (0.02% and 0.03% respectively) and may be fairly representative of the target population.

The ground truth of the test is shown in Appendix 6. The distribution of scores was approximately normal, shown in Figure 4, and average correct calls and miscalls for all signatures per group are shown in Figure 5.

Overall, no participant achieved a score below 20%, and the highest score attained was 90%, by three FDEs (1 fluent and 2 non-fluent). When the confidence of giving an opinion is factored into the results, the average correct calls per group are lower, as shown in Figure 6, and scores of 0% and 100% were obtained in a few instances. This was most noticeable in the FDEs, who showed more caution in their opinions, even though they were correct, than the laypeople, especially the fluent laypeople.

The questioned signatures which generated the greatest number of erroneous scores were Q3 and Q4. The results, detailed in full in Appendix 7, may be summarised as follows.

Analysis 1 shows that there was no difference between the ability of an FDE who is not fluent in Chinese and reports on signatures in Chinese, and the ability of an FDE not fluent in Chinese who does not report on signatures in Chinese, to assess authorship of a signature in Chinese script, irrespective of where the threshold of correct vs. incorrect answers was set (average correct scores 62.7% and 63.3% respectively; combined average 63%).

There was also no difference between the ability of an FDE who is fluent in Chinese, and the ability of an FDE not fluent in Chinese, to assess authorship of a signature in Chinese script (average correct scores 66% and 63% respectively). In other words, non-fluent FDEs who in real life do not report on Chinese signatures are just as accurate to those who do, and fluency does not help in the ability to assess authorship.

As expected, in general there was a significant difference between the ability of an FDE who is fluent in Chinese, and the ability of a layperson

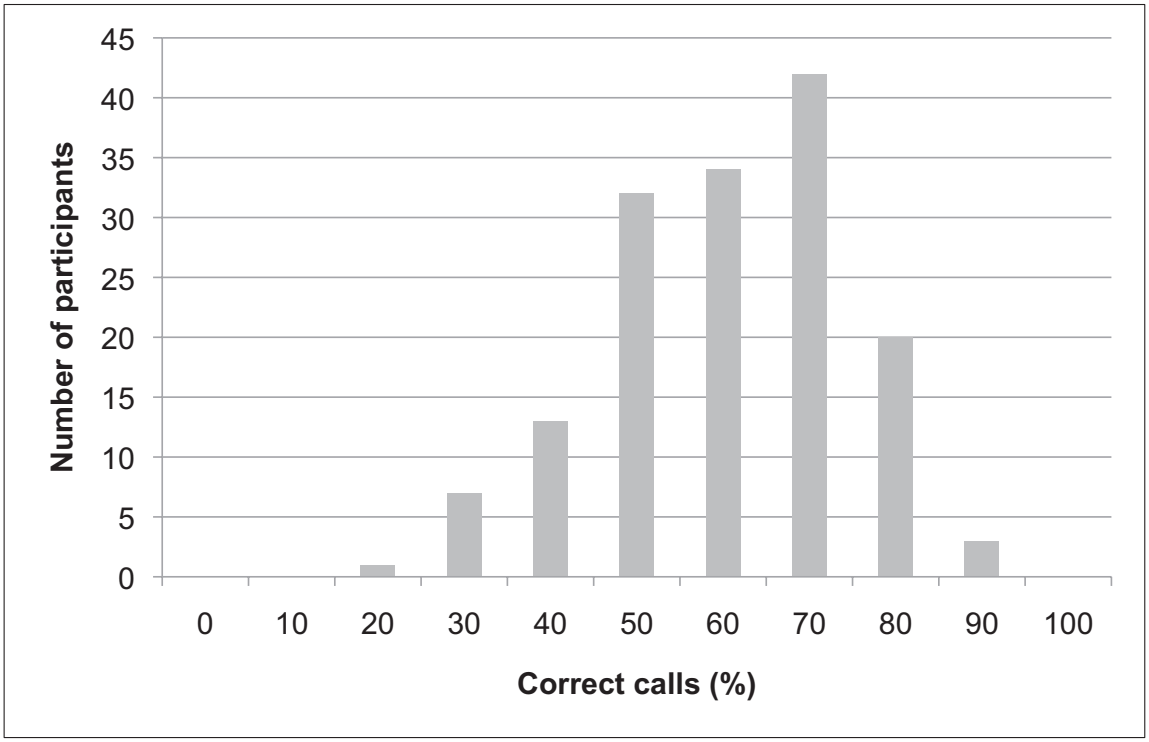

Figure 4. Distribution of all scores.

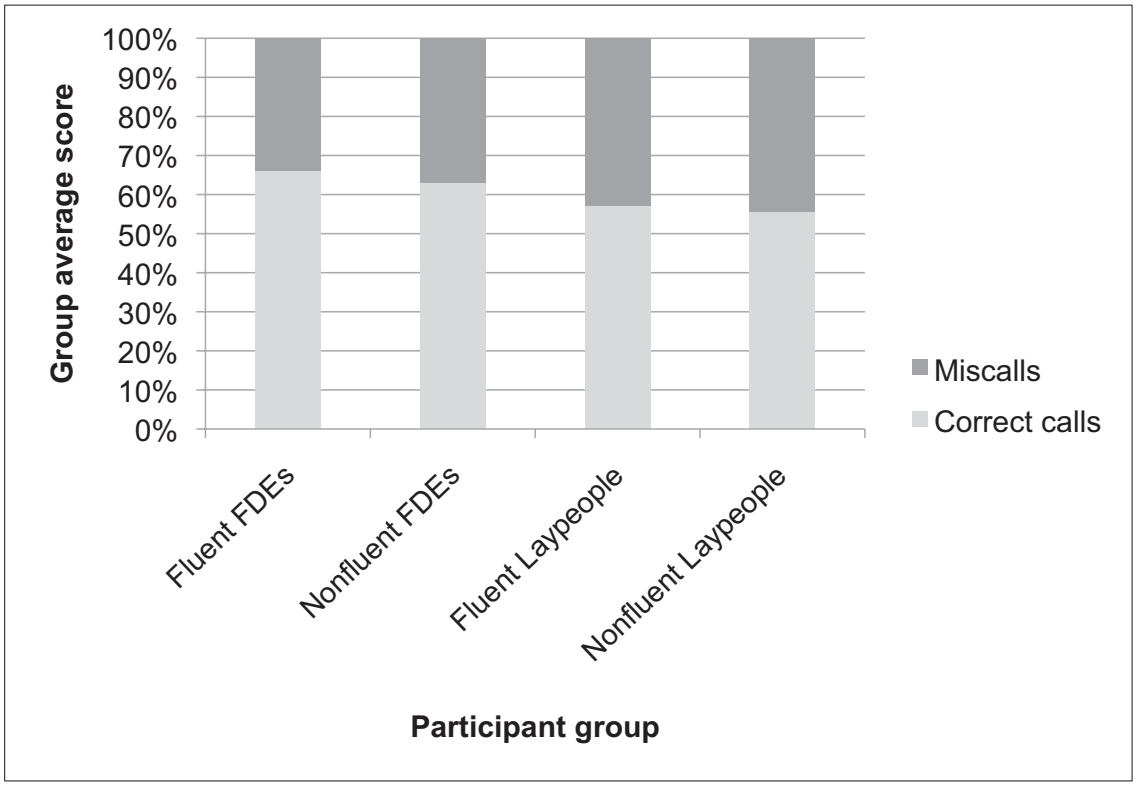

Figure 5. Average correct calls and miscalls for all signatures, by participant group.

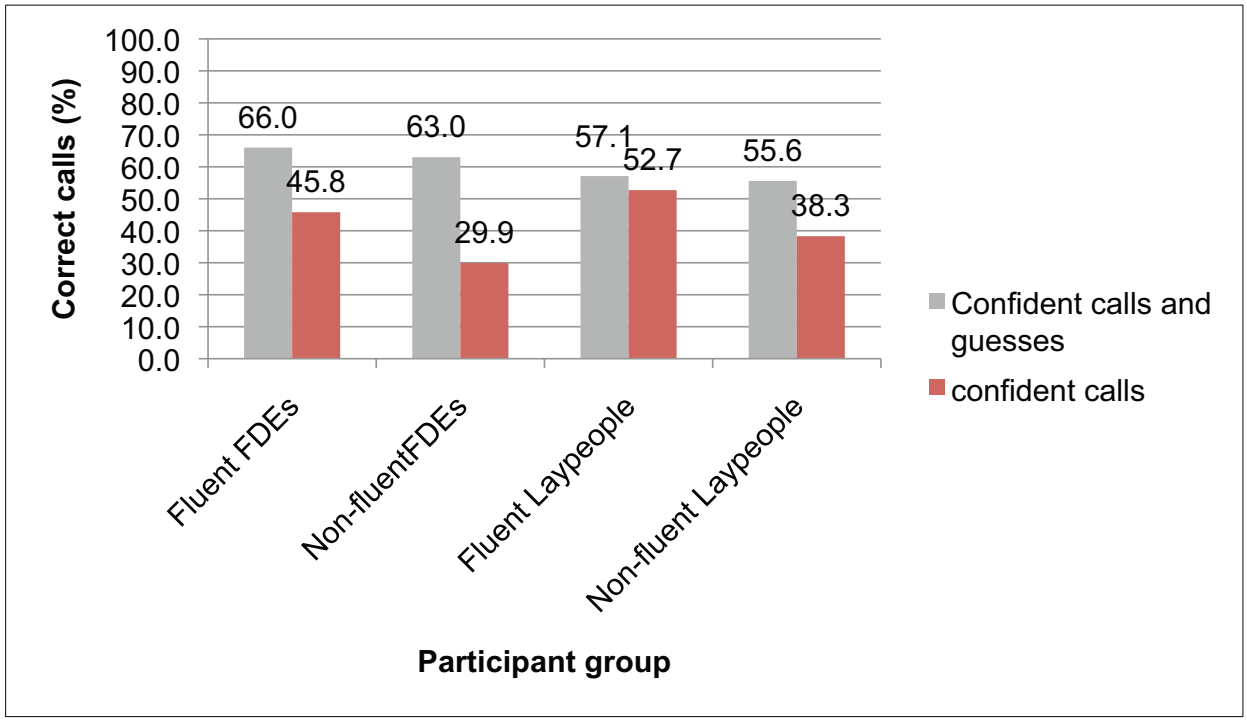

Figure 6. Comparison between all correct calls and only definite correct calls.

fluent in Chinese, to assess authorship of a signature in Chinese script (average correct scores 65.4% and 57.1% respectively). There was also a difference between the ability of an FDE who is not fluent in Chinese, and the ability of a layperson not fluent in Chinese, to assess authorship of a signature in Chinese script (average correct scores 63% and 55.6% respectively). However when the threshold was set quite high (7 or more correct answers), the ability of an FDE who is not fluent in Chinese, and the ability of a layperson not fluent in Chinese were the *same*. The main reason for this apparent anomaly was probably the small number of non-fluent FDEs who took part in the study. The numbers of FDEs and laypeople who achieved this were 10 and 7 respectively.

Analysis 2 accounted for repeated responses or measurements, and analysed everything together as a whole. It shows that ultimately fluency in a language *does not factor* the ability to assess authorship of a signature in Chinese script, but whether or not one is an FDE, *does*. The odds of correctly authenticating a questioned signature were 34% greater for the group of FDEs (in general) than for the group of laypersons. The analysis also showed that the choice of questioned signature being assessed has an effect on correct authentication. In other words, the accuracy of assessment depends on the individual signature in question.

Discussion

This study was not a simple test for FDEs. The many subtle complexities in the design, although often encountered in casework, are not usually incorporated into proficiency testing and other trials. Questioned signatures 3 and 4, from the ambidextrous writer, presented problems for the vast majority of respondents. This is most likely because the known specimens were written with the right hand and the questioned, although genuine, were written with the left. Here is a reminder of the need for comparable specimens, and of the possibility of encountering disguise. Whilst not an ideal comparison, this combination of known and questioned was generated by the random sampling method, and left in the study to see what the result would

be. Had these two questioned signatures been removed from the results, the proportion of correct scores would have been much higher. Also, the two questioned signatures, whilst fluently written, represented the greatest deviation, in terms of character construction, direction and shape of strokes, from the right-handed samples provided, and had a different pair of left-handed specimens been chosen for the trial, the results may have been quite different.

Successful forgers must be able to suppress their own writing habits and at the same time assume those unfamiliar characteristics of the original writer. The result of simulation is also related to writing skill (Leung[2] et al., 1993). This study used adult forgers who were completely at home with Chinese script, having learnt it as their first language, and who came from the highly-educated class of China. They were offered no monetary incentive – their enthusiasm stemmed from curiosity and satisfaction in taking part in what was probably for them a unique kind of research study. This had the potential to generate a pool of high quality simulations in the test samples with, one would suppose, a reasonable chance of fooling laypeople and FDEs alike, given the limitations of the test environment for conventional FDE evaluations and comparisons. The results appear to bear this out.

In this study, FDEs were more inclined to offer opinions with reduced confidence than laypeople. This correlates with other studies where laypeople tended to grossly over-associate (mismatch) unknown documents (Kam *et al.*, 1997). Laypeople seldom used the 'I'm not sure/educated guess' category for their answers, whereas FDEs did so frequently. In areas of science which involve a degree of quantitative analysis the term *measurement uncertainty* is commonly used. By offering a qualified rather than a definite conclusion, an examiner is merely expressing the greater uncertainty on which their opinion is based. In other words, there is less confidence because there is more uncertainty. Using this study as an example, the level of uncertainty noted by the FDEs would probably have decreased had they been provided with original samples, or more comparable exemplars, and this is what prevented them from rendering a more definite conclusion. More uncertainty does not equate to a greater probability of error and the two concepts are not equivalent.

Overall, performance of the FDEs was superior to the performance of laypersons, correlating with previous research studies (Kam *et al.*, 1997, 2001, 2003). The fluent FDEs scored more highly than the non-fluent FDEs, as might be expected, which may be explained by their greater familiarity with the script. It is possible that the scores for fluent FDEs might have been even higher had they possessed more years of experience, since a fairly large proportion of them had only between 1 and 2 years' experience. Of the non-fluent FDEs, those who do currently express opinions on foreign script signatures (and no information was presented as to which scripts these might be, although Chinese may have been inadvertently implied), had similar scores to those non-fluent FDEs who do not currently express opinions on signatures written in a foreign script. These preliminary results suggest that the skills possessed by a properly-qualified FDE are sufficient for the identification of signatures written in Chinese script, and a perceived or true lack of experience or confidence may not be a detrimental factor in the accuracy of the results. These results align with those expressed by Al-Musa Alkahtani, relating to the ability of FDEs to successfully recognize forgeries in foreign signatures with which they are not familiar, namely Arabic script (Al-Musa Alkahtani, 2010, 2014). The second study further supported the belief that literacy in a script is not needed to detect forgery in that script, but also noted that other considerations may mean that a native writer of a script may still have advantages over an illiterate in that script in conducting forensic handwriting examination. Examples of these include the invention and use of abstract designs commonly found in calligraphic script cultures, such as Arabic and Chinese. Common customs such as these, and their appropriateness, might only be understood by a script-literate individual.

The finding that whilst some non-fluent FDEs have not reported on Chinese signatures, they would in fact do so with similar accuracy to those who do is encouraging but should be treated with caution. It must be emphasised that in this study there was a low study sample (only 26 formally-trained non-fluent FDEs took part in the study, and these were further divided into those who don't report and those who do report, on foreign signatures). Thus the findings should not be extrapolated to the overall FDE population. It would also be interesting to explore the reasons why some non-fluent FDEs don't report on 'foreign' signatures, whilst others do.

No peer review was allowed in this study, but many participants stated that their usual practice

was to always employ peer review. With peer review the FDE scores may have been much higher had this been allowed, since although peer review can reduce the call rate, in this study inconclusive answers were not allowed. In future this should be permitted, not only for its desirability in FDE examinations, but to reflect real life more accurately.

Another study, focusing on whether individual signatures were genuine, is reported in Sita *et al*. (2002). It was found that professional handwriting examiners erred in 3.4% of their judgments, but inconclusive responses were excluded from this statistic. The error rate in the present study was higher due to the forced-call protocol. De facto this removes the call on adequacy i.e. deciding at the outset whether there is sufficient material for a comparison to be made. Professional FDEs are trained to take this into consideration whilst laypeople tend to overlook it (Ostrum, 2013).

Questioned signatures Q6, Q7, and Q9 were written in a somewhat constrained space on the consent form (Figure 5) which reduced their size, and may have affected their style and proportions, added or removed radicals, or induced other artefacts, as has been shown to occur in Latin-script signatures (Dainis *et al*., 2011). It is possible that this form may have been regarded by the writers as a more formal document than the request signature sheets, and thus a more careful signature was employed. The 12pt font used for the printed text on the form may also have induced cognitive or contextual bias, leading participants to reduce the size of their signature to align more with the surrounding text. This form was always signed prior to any request signature specimens.

Study feedback from participants

Various comments from those who took part in this study were received. A number of FDEs in China commented on the quality of the simulations, viewing the study as a serious undertaking, and were very interested to know the ground truth. Several of the non-fluent FDEs expressed concern that the known specimens were requested specimens, not course-of-business writings and thus inadequate for comparison purposes.

Another concern was that of using non-original writings, which has been raised previously with similar large-scale trials (ST2AR[6]), since reproductions can never show all the fine detail upon which an FDE relies, in order to give an opinion. This is a shortcoming which is difficult to surmount although studies using original and non-

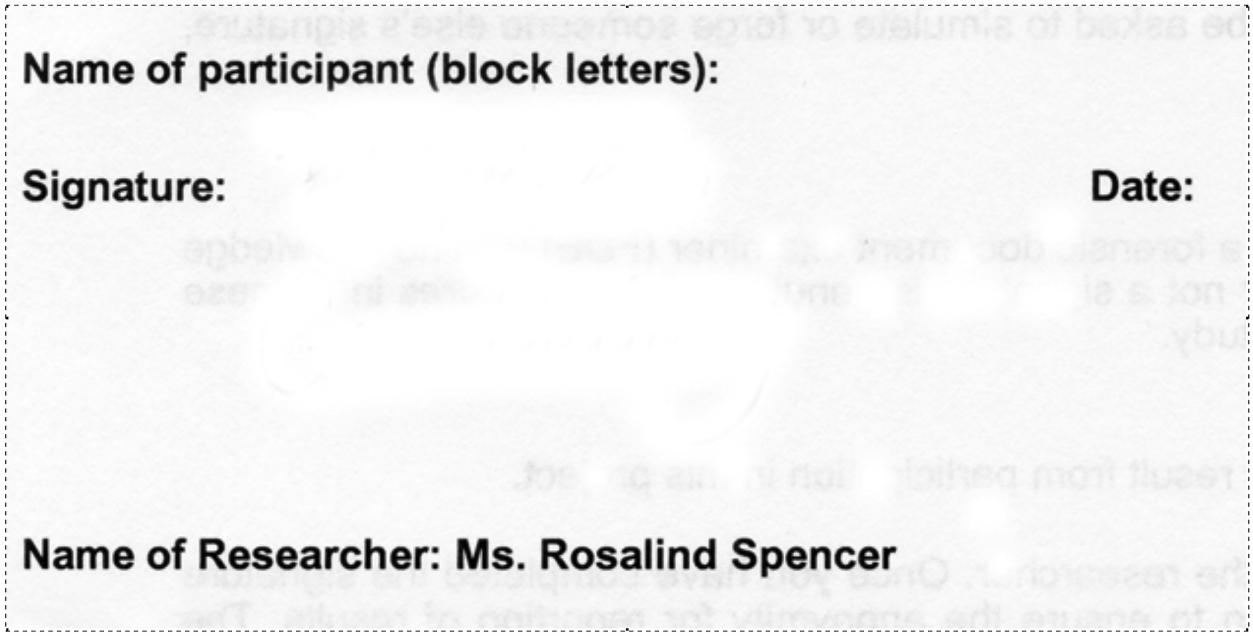

Figure 5. Constrained area for signing on consent form (Q6, Q7 and Q9) (not to scale, border was not delineated on form).

[6]Skill-TaskTraining Assessment & Research, Inc. www.ST2AR.org.

original material have not borne this out. Digital images captured by scanning often provide useful information about the handwriting and cannot be disregarded. In some instances, scanned images, especially those scanned in color and at higher-resolutions, are sufficient for a reliable evaluation of the handwriting (Masson, 2012). Dawson and Lindblom (1998) also demonstrated that accurate judgements can be made using non-original material, in that forensic document examiners miss very few elements of line quality when presented with average quality first generation photocopies.

Another remarked that all the simulations were made by well-educated, fluent forgers, with no comparison for either poorly educated or non-fluent forgers. Finally, on the participant questionnaire one layperson (a schoolteacher for special-needs children) misinterpreted the question relating to 'Do you sign in English or Chinese' with signing language for the deaf community, highlighting the importance of unambiguous terminology in questionnaires!

The study has the potential to be adapted as a model and applied to signatures in other non-Latin scripts. It is hoped that in the future further research studies in such scripts will follow.

Acknowledgements

The author is indebted to all the participants worldwide for kindly agreeing to take part in this study. Particular thanks go to Professor Linda Liu, Jiangsu Police Institute, China for her assistance with coordinating the project in mainland China. Assistance with initial experimental design and the subsequent statistical analysis of the results was provided by Sofia Auer, Mathematician, CBSA.

References

Al-Musa Alkahtani, A (2010) The Ability of Forensic Handwriting Examiners to Judge the Quality of Signature Simulations in an Unfamiliar Writing System. JASQDE, 13 No 2.

Al-Musa Alkahtani, A (2014) Arabic-Illiterate Forensic Handwriting Analysis: A Pilot Study to Further Investigate the Ability of Arabic-Illiterate Examiners to Judge the Accuracy of Simulations of Arabic Signatures, Compared with Arabic-Literate Examiners. JASQDE, 17 No 1.

Brewester F (1932) Contested Documents and Forgeries. The Book Company Ltd, Calcutta.

Cheng Y (2011) An overview of Radicals in Chinese Characters Part II—A study on the writing of radicals of similar structures. Proceedings of the ASQDE Annual Conference, Victoria BC.

Coulmas F (1999) The Blackwell Encyclopedia of Writing Systems. Blackwell Publishing)

Crown D and Shimaoka T (1974) The Examination of Ideographic Handwriting. Journal of Police Science and Administration, Northwestern University School of Law.

Dawson GA and Lindblom BS (1998) An evaluation of line quality in photocopied signatures. Science & Justice 38(3) 189-194.

Desjardins R (2012) The Effects of Constraint on the Static and Dynamic Features of Handwriting. University of Ontario Institute of Technology B.Sc (Hons) thesis.

Dewhurst T, Found B and Rogers D (2008) Are expert penmen better than lay people at producing simulations of a model signature? Forensic Sciences International 180 (1) 50-53.

Fazio K (2011) The Effects of Constraint on a Signature's Static and Dynamic Features. University of Ontario Institute of Technology B.Sc (Hons) thesis.

Johnston Y (2012) Understanding the Chinese Writing System 101. Proceedings of the ASQDE Annual Conference, Charleston.

Kam M, Fielding G & Conn R (1997) Writer identification by professional document examiners. Journal of Forensic Sciences 42 (5) 778-786.

Kam M, Gummadidala M, Fielding G and Conn R (2001) Signature authentication by forensic document examiners. Journal of Forensic Sciences 46 (4) 884-888.

Kam M, Gorski A and Gaughan C (2003) A Decade of Writer Identification Proficiency Tests for Forensic Document Examiners. Proceedings of the ASQDE Annual Conference, Baltimore.

Leung[1] S, Cheng Y, Fung H and Poon N (1993) Forgery I - Simulation; Journal of Forensic Sciences, Vol. 38, No. 2: 402-412.

Leung[2] S, Fung H, Cheng Y and Poon N (1993) Forgery II - Tracing; Journal of Forensic Sciences, Vol. 38, No. 2:413-424.

Masson J F (2012) SCANNED IMAGES: How Well Do They Depict the Subtle features in Handwriting? JASQDE, 15 No 1.

Mittal S & Sehgal V (1989) Forensic Examination of Unfamiliar Scripts, International Criminal Police Review, No. 418, pp 11-15.

National Academy of Sciences, National Research Council, Committee on Identifying the Needs of the Forensic Science Community. Strengthening Forensic Science in the United States: A Path Forward. National Academies Press, 2009.

Ostrum, R (2013) Error, Confidence and (Un) certainty—Deconstructing Authorship Opinions Using a Forced Call Testing Protocol. Proceedings of National Institute of Standards and Technology Measurement Science and Standards in Forensic Handwriting Analysis Conference, Gaithersburg, ML.

Simsons D, Spencer R and Auer S (2011) The Effects of Constraining Signatures. Journal of the American Society of Questioned Document Examiners Vol. 14, no.1.

Sita J, Found B and Rogers D (2002) Forensic handwriting examiners' expertise for signature comparison. Journal of Forensic Sciences 47:1117.

Wilder D and Ingram J (1974) Analysis of Chinese Characters, Dover Publications Inc.

Appendix 1. Participant Questionnaire

PARTICIPANT QUESTIONNAIRE

Project title: Investigating forensic document examiner opinions on signatures in a foreign script

Researcher: Rosalind Spencer

To be completed by the study participant:

Please check/circle the appropriate answer

1 Sex: Male Female

2. Age: 18-30 31-50 51+

3. Highest level of education:
High school
University
Other (please specify) _____

4. Do you consider yourself fluent(able to easily identify, read, write, sign) in English script?
Yes No

5. Do you consider yourself fluent(able to easily identify, read, write, sign) in Chinese script?
Yes No

6. In which country did you learn English script (if applicable)? _____

7. In which country did you learn Chinese script (if applicable)? _____

8. What do you consider is your first language? _____

9. In what language(s) do you usually sign? _____

10. I am: Right-handed Left-handed

11. On average how many times per week do you sign your name?
0-9
More than 10

12. Have you taken any courses in Chinese calligraphy? Yes No

13. If you are not a Forensic Document Examiner, what is your main area of study/expertise/work?

All answers will be treated in the strictest confidence and will be made anonymous
所有的回答将绝对保密地对待，匿名处理。

If you are a Forensic Document Examiner, please complete Additional Survey (Page 3).

I understand that I can request results of this project by email when it is completed, from Ms. Rosalind Spencer (Rosalind.spencer@cbsa-asfc.gc.ca).

Additional Survey on Rendering Opinions on Signatures Written in Chinese Characters
中文签名检验意见的其他调查
All answers will be treated in the strictest confidence and will be made anonymous
所有的回答将绝对保密地对待，匿名处理。

1. How many years of forensic document examination experience do you have?_____

2. Which statement best describes your experience of Chinese characters, Chinese signatures and the Chinese language? Fluency = ability to easily identify, read, write, and sign in Chinese script

   l am fluent in Chinese

   I am not fluent but I have some knowledge of Chinese script

   I have no knowledge of Chinese script

3. If you answered 'I am not fluent but I have some knowledge of Chinese script', how was it gained

4. Do you give opinions on signatures written in Chinese characters? (other than 'inconclusive')?

   Yes No

5. On how many Chinese signatures a year do you give such an opinion _____

6. What identifying or eliminating opinions have you given on a signature(s) written in Chinese characters? (circle all those closest to the opinions you have rendered)

   Identification
   Strong probability did write
   Probably did write
   Some evidence did write
   Probably did not write
   Strong probability did not write
   Elimination

7. If you had to describe your personal approach to forensic signature examination in general, which of the following would best describe it?

   - A 'class/individual' paradigm (you believe your expertise is based on knowing what is class and what is individual and it is your extensive knowledge of how individual features are that enables you to identify)
   - A 'complexity' paradigm (you make assumptions about the variation in the population of features to enable you to identify).
   - Not sure.

Thank you very much for your participation.

Any other comments you may wish to convey

Appendix 2. Sheet used for request specimens and simulations

Appendix 3. Instruction sheet

INSTRUCTIONS

Please note: Signatures used in this study contain person-identifiable information, such as a legible signature. As part of ongoing research conducted at CBSA, participants consented to the use of their signature for research with the promise that their identities would not be inappropriately disclosed. Please respect the confidentiality of the person by not sharing identifiable information with others outside the study group.

1. Please fill out the Participant Questionnaire attached.
2. You have been provided with sets of signatures in 5 different names (printed and digital). Each set comprises 2 questioned signatures, and 12 known specimens.
3. You may use any combination of these digital or printed images to aid your examination.
4. All genuine signatures and known specimens were written in a single sitting. The gridlines you see on the knowns were present at the time but without any text. Assume that they are all fluently and naturally written. They are contemporaneous with the questioned signatures and are life-size 1:1.
5. All questioned signatures were written either within similar gridlines as described above or within a similar space surrounded by text. The text has been removed to reduce contextual bias. The questioned signatures are independent of each other (i.e. not related) and are life-size 1:1. Please treat them separately.
6. Decide whether or not the specimen writer wrote the questioned signatures.
7. On the Answer Sheet provided, please circle your answer in the Answer column.
8. If you answered 'don't know', please indicate what your educated guess would be, in the next column.
9. Any comments are most welcome and may be continued on a separate sheet.
10. Please do not discuss your answers with anyone.
11. Please do not conduct any peer review with another examiner.
12. To save resources some of you may be asked to share this package with another examiner, so please maintain the integrity of the printed images and electronic images. After final use, please destroy the images.
13. If you are sharing study materials with another examiner, please make additional copies of the participant questionnaire and answer sheet as necessary.

Completed questionnaires and answer sheets may be forwarded by email, post or fax.

Appendix 4. Participant answer sheet

Investigating Forensic Document Examiner Opinions on Signatures in a Foreign Script
ANSWER SHEET

Signature		Answer	I don't know / educated guess	Please give comments here or on a separate sheet
2	Q1	It was Writer 2 It was not Writer 2 I don't know	It was Writer 2 It was not Writer 2	
2	Q2	It was Writer 2 It was not Writer 2 I don't know	It was Writer 2 It was not Writer 2	
4	Q1	It was Writer 4 It was not Writer 4 I don't know	It was Writer 4 It was not Writer 4	
4	Q2	It was Writer 4 It was not Writer 4 I don't know	It was Writer 4 It was not Writer 4	
6	Q1	It was Writer 6 It was not Writer 6 I don't know	It was Writer 6 It was not Writer 6	
6	Q2	It was Writer 6 It was not Writer 6 I don't know	It was Writer 6 It was not Writer 6	
7	Q1	It was Writer 7 It was not Writer 7 I don't know	It was Writer 7 It was not Writer 7	
7	Q2	It was Writer 7 It was not Writer 7 I don't know	It was Writer 7 It was not Writer 7	
10	Q1	It was Writer 10 It was not Writer 10 I don't know	It was Writer 10 It was not Writer 10	
10	Q2	It was Writer 10 It was not Writer 10 I don't know	It was Writer 10 It was not Writer 10	

Appendix 5. Portrait of the participants

Group	Number in group	Geographical region
Fluent FDEs	48	China, Hong Kong, Singapore
Non-fluent FDEs Do report Don't report	 11 15	Canada, USA, UK, Continental Europe, Australia
Fluent laypeople	62	China, Canada
Non-fluent laypeople	18	Canada, USA, UK
Total participants	154	

Appendix 6 Test sample set ground truth

Test specimen	Study ground truth	
	Known signatures	*Questioned signature*
Q1	Writer 2 request specimens written with right hand	Simulation by writer 5
Q2	Writer 2 request specimens written with right hand	Simulation by writer 5
Q3	Writer 4 request specimens written with right hand	Genuine signature on request specimen form, written with left hand
Q4	Writer 4 request specimens written with right hand	Genuine signature on request specimen form, written with left hand
Q5	Writer 6 request specimens written with right hand	Simulation by writer 10
Q6	Writer 6 request specimens written with right hand	Genuine signature on consent form
Q7	Writer 7 request specimens written with right hand	Genuine signature on consent form
Q8	Writer 7 request specimens written with right hand	Simulation by writer 9
Q9	Writer 10 request specimens written with right hand	Genuine signature on consent form
Q10	Writer 10 request specimens written with right hand	Simulation by writer 6

Appendix 7: Analysis of Results

Analysis 1: Chi-Squared Testing

Hypothesis 1

H_0: The distribution of the variable 'correct assessment of authenticity' is the same across populations (where population one includes FDEs fluent in Chinese and population two includes FDEs not fluent in Chinese).

H_1: The distribution of the variable 'correct assessment of authenticity' differs across populations (where population one includes FDEs fluent in Chinese and population two includes FDEs not fluent in Chinese).

Method

- Each participant in the study was given ten signatures for assessment of authenticity. These ten signatures included five writers' signatures, each repeated twice.
- These ten responses per participant, were then used to create a new categorical variable "Correctly Assessed" with two levels (High, Low); a participant would be categorized under "High" if they were able to correctly assess between 6 and 10 signatures, and "Low" if they were able to correctly assess between 0-5 signatures.
- Testing was done three times, corresponding to three different threshold splits of the variable "Correctly Assessed" (i.e. second test used a"High"/ "Low" split of 5-10 signatures (High) and 0-4 signatures (Low), and third test used a "High"/ "Low" split of 7-10 signatures (High) and 0-6 signatures (Low)).

Analysis

Three chi-squared tests were performed to test the null hypothesis. Each test corresponded to a different threshold split of the response variable as per description in the Methods section above.

Results

Corresponding to the three different response variable threshold splits:

(1) Threshold: Low (0-5), High (6-10)

The distribution of correct authenticity assessments is the same between the population of FDEs fluent in Chinese and the population of FDEs not fluent in Chinese, $X^2(1, N=74)=0.0002$, $p=0.9875$.

(2) Threshold: Low (0-4), High (5-10)

The distribution of correct authenticity assessments is the same between the population of FDEs fluent in Chinese and the population of FDEs not fluent in Chinese, $X^2(1, N=74)=0.0045$, $p=0.9468$.

(3) Threshold: Low (0-6), High (7-10)

The distribution of correct authenticity assessments is the same between the population of FDEs fluent in Chinese and the population of FDEs not fluent in Chinese, $X^2(1, N=74)=2.1346$, $p=0.144$.

Hypothesis 2

H_0: The distribution of the variable 'correct assessment of authenticity' is the same across populations (where population one includes FDEs who are not fluent in Chinese and report on signatures in Chinese script, and population two includes FDEs who are not fluent in Chinese and do not report on signatures in Chinese script).

H_1: The distribution of the variable 'correct assessment of authenticity' differs across populations (where population one includes FDEs who are not fluent in Chinese and report on signatures in Chinese script, and population two includes FDEs who are not fluent in Chinese and do not report on signatures in Chinese script).

Method

- Same as for Hypothesis 1.

Analysis

Two chi-squared tests were performed to test the null hypothesis. Each test corresponded to a different threshold split of the response variable as per description in the Methods section above.

Results

Corresponding to the three different response variable threshold splits:

(1) Threshold: Low (0-5), High (6-10)

The distribution of correct authenticity assessments is the same between the population of FDEs who are not fluent in Chinese and **report** on signatures in Chinese script, and the population of FDEs who are not fluent in Chinese and **do not report** on signatures in Chinese script, $X^2(1, N=26)=0.1891$, $p=0.6637$.

(2) Threshold: Low (0-4), High (5-10)

Insufficient sample size for testing (i.e. 50% of Chi-squared grid cells had <5 counts)

(3) Threshold: Low (0-6), High (7-10)

The distribution of correct authenticity assessments is the same between the population of FDEs who are not fluent in Chinese and **report** on signatures in Chinese script, and the population of FDEs who are not fluent in Chinese and **do not report** on signatures in Chinese script, $X^2(1, N=26)=0.0355$, $p=0.8506$.

Hypothesis 3

H_0: The distribution of the variable 'correct assessment of authenticity' is the same across populations (where population one includes FDEs fluent in Chinese and population two includes laypersons fluent in Chinese).

H_1: The distribution of the variable 'correct assessment of authenticity' differs across populations (where population one includes FDEs fluent in Chinese and population two includes laypersons fluent in Chinese).

Method

- Same as for Hypothesis 1.

Analysis

Three chi-squared tests were performed to test the null hypothesis. Each test corresponded to a different threshold split of the response variable as per description in the Methods section above.

Results

Corresponding to the three different response variable threshold splits:

(1) Threshold: Low (0-5), High (6-10)

The distribution of correct authenticity assessments differs between the population of FDEs fluent in Chinese and the population of laypersons fluent in Chinese, at an alpha level of $\alpha=0.05$, $X^2(1, N=110)=4.384$, $p=0.03628$.

OR

A higher proportion of the FDEs fluent in Chinese (77%, or 37 of 48) correctly authenticated a higher number of signatures than did the group of laypersons fluent in Chinese (58%, or 36 of 62), $X^2(1, N=110)=4.384$, $p=0.03628$.

OR

The odds of correctly authenticating a high number of questioned signatures were 2.43 times greater for FDEs fluent in Chinese than for laypersons fluent in Chinese (OR=2.43, 95% CI=1.05, 5.63). Thus, the odds ratio is greater than 1, suggesting that the true odds of correctly authenticating a high number of questioned signatures are greater for FDEs fluent in Chinese.

(2) Threshold: Low (0-4), High (5-10)

The distribution of correct authenticity assessments differs between the population of FDEs fluent in Chinese and the population of laypersons fluent in Chinese, at an alpha level of $\alpha=0.05$, $X^2(1, N=110)=7.3803$, $p=0.0066$.

OR

A higher proportion of the FDEs fluent in Chinese (96%, or 46 of 48) correctly authenticated a higher number of signatures than did the group of laypersons fluent in Chinese (77%, or 48 of 62), $X^2(1, N=110)=7.3803$, $p=0.0066$.

OR

The odds of correctly authenticating a high number of questioned signatures were 6.71 times greater for FDEs fluent in Chinese than for laypersons fluent in Chinese (OR=6.71, 95% CI=1.44, 31.16). Thus, the odds ratio is significantly greater than 1, suggesting that the true odds of correctly authenticating a high number of questioned signatures are greater for FDEs fluent in Chinese.

(3) Threshold: Low (0-6), High (7-10)

The distribution of correct authenticity assessments differs between the population of FDEs fluent in Chinese and the population of laypersons fluent in Chinese, at an alpha level of $\alpha=0.05$, $X^2(1, N=110)=5.509$, $p=0.0189$.

OR

A higher proportion of the FDEs fluent in Chinese (56%, or 27 of 48) correctly authenticated a higher number of documents than did the group of laypersons fluent in Chinese (34%, or 21 of 62), $X^2(1, N=110)=5.509$, $p=0.0189$.

OR

The odds of correctly authenticating a high number of questioned signatures were 2.51 times greater for FDEs fluent in Chinese than for laypersons fluent in Chinese (OR=2.51, 95% CI=1.16, 5.45). Thus, the odds ratio is greater than 1, suggesting that the true odds of correctly authenticating a high number of questioned signatures are greater for FDEs fluent in Chinese.

Hypothesis 4

H_0: The distribution of the variable 'correct assessment of authenticity' is the same across populations (where population one includes FDEs not fluent in Chinese and population two includes laypersons not fluent in Chinese).

H_1: The distribution of the variable 'correct assessment of authenticity' differs across populations (where population one includes FDEs not fluent in Chinese and population two includes laypersons not fluent in Chinese).

Method

- Same as for Hypothesis 1.

Analysis

Three chi-squared tests were performed to test the null hypothesis. Each test corresponded to a different threshold split of the response variable as per description in the Methods section above.

Results

Corresponding to the three different response variable threshold splits:

(1) Threshold: Low (0-5), High (6-10)

The distribution of correct authenticity assessments differs between the population of FDEs not fluent in Chinese and the population of laypersons not fluent in Chinese, at an alpha level of $\alpha=0.05$, $X^2(1, N=44)=4.8486$, $p=0.02767$.

OR

A higher proportion of the FDEs not fluent in Chinese (77%, or 20 of 26) correctly authenticated a higher number of signatures than did the group of laypersons not fluent in Chinese (44%, or 8 of 18), $X^2(1, N=44)=4.8486$, $p=0.02767$.

OR

The odds of correctly authenticating a high number of questioned signatures were 4.17 times greater for FDEs not fluent in Chinese than for laypersons not fluent in Chinese (OR=4.17, 95% CI=1.13, 15.33). Thus, the odds ratio is significantly greater than 1, suggesting that the true odds of correctly authenticating a high number of questioned signatures are greater for FDEs not fluent in Chinese.

(2) Threshold: Low (0-4), High (5-10)

The distribution of correct authenticity assessments differs between the population of FDEs not fluent in Chinese and the population of laypersons not fluent in Chinese, at an alpha level of $\alpha=0.10$, $X^2(1, N=44)=3.5659$, $p=0.05898$.

OR

A higher proportion of the FDEs not fluent in Chinese (96%, or 25 of 26) correctly authenticated a higher number of signatures than did the group of laypersons not fluent in Chinese (78%, or 14 of 18), $X^2(1, N=44)=3.5659$, $p=0.05898$.

OR

The odds of correctly authenticating a high number of questioned signatures were 7.14 times greater for FDEs not fluent in Chinese than for laypersons not fluent in Chinese (OR=7.14, 90% CI=1.05, 48.69). Thus, the odds ratio is significantly greater than 1, suggesting that the true odds of correctly authenticating a high number of questioned signatures are greater for FDEs not fluent in Chinese.

(3) Threshold: Low (0-6), High (7-10)

The distribution of correct authenticity assessments **is the same** between the population of FDEs not fluent in Chinese and the population of laypersons not fluent in Chinese, at an alpha level of $\alpha=0.05$, $X^2(1, N=44)=0.0008$, $p=0.9772$.

Analysis 2: Generalized Estimating Equations for Modeling Repeated Binary Categorical Outcome Variables

Method

Generalized estimating equations (GEE) were used to model repeated binary response variable (correct authentication of a questioned signature). The predictor variables used included language fluency, FDE status, and questioned signature written in a particular name.

Results

Tables 1 and 2 below indicate that language fluency is not a factor in correctly authenticating a questioned signature and neither is the interaction between language fluency and FDE status. However, whether or not a participant was an FDE did have a significant association ($p=0.0441$), as did the particular questioned signature being assessed ($p<0.0001$).

Table 1. Score statistics for type 3 GEE analysis

Score Statistics For Type 3 GEE Analysis			
Source	**DF**	**Chi-Square**	**Pr > ChiSq**
Language fluency	1	0.01	0.9312
FDE_status	1	4.05	0.0441
Signature in a particular name	1	18.83	<.0001
fluency*FDE_status	1	0.03	0.8589

Table 2. Analysis of GEE parameter estimates

Analysis Of GEE Parameter Estimates								
Empirical Standard Error Estimates								
Parameter			Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept			-0.2689	0.2269	-0.7135	0.1758	-1.19	0.2360
Language fluency (fluent)	1		-0.0126	0.2173	-0.4386	0.4134	-0.06	0.9537
Language fluency (non-fluent)	0		0.0000	0.0000	0.0000	0.0000	.	.
FDE_status (FDE)	1		0.2650	0.2383	-0.2020	0.7320	1.11	0.2660
FDE_status (non-FDE)	0		0.0000	0.0000	0.0000	0.0000	.	.
Signature in a particular name			0.1013	0.0225	0.0572	0.1454	4.51	<.0001
Language fluency *FDE_status	1	1	0.0491	0.2767	-0.4933	0.5914	0.18	0.8592
Language fluency *FDE_status	1	0	0.0000	0.0000	0.0000	0.0000	.	.
Language fluency *FDE_status	0	1	0.0000	0.0000	0.0000	0.0000	.	.
Language fluency *FDE_status	0	0	0.0000	0.0000	0.0000	0.0000	.	.

Table 3 gives the point estimate for the odds ratio of FDE status, as OR= 1.3358 with 95% confidence limits of (1.01861, 1.7518). Thus the odds of correctly authenticating a questioned signature was 1.34 times greater for the group of FDEs (in general) than for the group of laypersons (in general), regardless of fluency in Chinese (interaction of language fluency and FDE status was tested and found to be non-significant, p=0.8589).

Table 3. Contrast estimates

Contrast Estimate Results										
Label	Mean Estimate	Mean Confidence Limits		L'Beta Estimate	Standard Error	Alpha	L'Beta Confidence Limits		Chi-Square	Pr > ChiSq
Language fluency	0.5030	0.4356	0.5703	0.0119	0.1383	0.05	-0.2592	0.2831	0.01	0.9313
Exp(fluency)				1.0120	0.1400	0.05	0.7716	1.3272		
status	0.5719	0.5046	0.6366	0.2895	0.1383	0.05	0.0185	0.5606	4.38	0.0363
Exp(status)				1.3358	0.1848	0.05	1.0186	1.7518		