Journal of the American Society of Questioned Document Examiners, Inc.
Volume 22, Number 1, June 2019

Forensic Investigative and Evaluative Assessment of Handwritten X-Marks

Raymond Marquis, Ph.D., Tacha Hicks, Ph.D., Williams Mazzella, Ph.D.

School of Criminal Justice, University of Lausanne, Switzerland

Analysis of simple written entries can represent a challenge in the practice of forensic handwriting examiners. This article shows that valuable information can be elicited from X-marks features, in agreement with the recommendations of the ENFSI guideline for evaluative reporting. In the course of a real case encountered by the authors, an experimental study was designed to acquire structured data required to help answer the questions of the mandate. X-marks from 75 right-handed and 25 left-handed writers were collected and classified according to their stroke sequence. The results of this empirical study were first used to assess handedness of the writer, together with a development on the risk of misleading evidence as a measure of the method performance. The results were then used to assess writership of a given person rather than an unknown person. This paper shows that following the ENFSI recommendations for evidence interpretation may require only a small dataset collected for the case needs. The procedure of evidence interpretation detailed in this paper may be followed by any examiner interested in applying a Bayesian approach on simple data collected for assessing the results of a given case, should this concern an X-mark or any other handwritten sign or letter.

Keywords: Handwriting, Signatures, X-marks, Interpretation, Likelihood ratio, Bayesian approach

Introduction

Analysis of simple written entries can represent a challenge in the practice of forensic handwriting examiners. In the field of signature examination, complexity is generally assumed to be a core element for the discrimination of genuine and forged signatures. It is considered that the more complex the signature, the more difficult it is to be simulated (Found and Rogers, 1998). In case of a simple signature, similarities between a questioned and reference signatures may be of limited value to help determine whether the questioned signature is genuine or forged, since similarities would be expected in both hypotheses.

In the study by Cadola *et al.* (2013), a set of signatures of six writers, considered to be simple, was sampled. These signatures were transmitted to 52 forgers, some of them being forensic students or presenting artistic skills, to produce free-hand simulations. The best simulations of three different signatures were submitted for evaluation to forensic handwriting examiners. None of the forged signatures was wrongly classified as a genuine signature.

Signature complexity is most commonly evaluated qualitatively, but models have been proposed to help examiners decide on the complexity level through quantification (Found and Rogers, 1998). Signature complexity seems to be strongly correlated to the number of crossings and direction changes. According to these models, an X-mark would be undoubtedly considered as very simple and classification mistakes could occur with such extreme simple signature.

On another hand, handwriting examiners may face a questioned X-mark that is not a signature, but a simple item of handwriting, such as a mark written in a box to validate a choice on a form. In the latter case, one may argue that such a handwriting item does not present enough variation between people to be of any value to point towards a given writer instead of any other. Nevertheless, as odd as it may seem, X-marks can embody information that can be of value in a case.

On the matter of X-marks, Osborn (1932) reported a testimony that was judged as early as 1927. At the time of this judgement, according to this author, testimony regarding the authenticity of X-marks was generally not admitted in many courts. However, according to Hilton (1982, p. 206), in spite of the apparent simplicity of such marks, many elements can be considered to help answer the question of writership of an X-mark. The position of the crossing point between the two strokes of the mark may vary between writers, as well as the orientation of these strokes, and whether these strokes are connected or not. The line quality of the strokes, related to the writing fluency, may also be helpful to discriminate between writers (Huber and Headrick, 1999, p.148). Other valuable features reside in the direction of the strokes and their sequence (Foley, 1999; Welch, 1999), as well as pen pressure and pen position (Hilton, 1982, p. 206). While the relative length of the strokes may also be useful in case where the X-mark is used as a signature (such as reported in Foley, 1999), it may not be relevant in cases where the marks are affixed in boxes, due to space constraints.

This article will present a case illustrating the value that can be elicited from X-marks, both in investigative and evaluative proceedings. After a brief description of the case and the request from the mandating authority, we will expose the results of an experimental study carried out for the case needs. These results will first be used to assess handedness of the writer, together with a development on the risk of misleading evidence as a measure of the method performance. They will then help assess writership of a given person rather than an unknown person.

The case

The case that motivated this article involves an X-mark affixed in a box on a questioned document. This mark was meant for the validation of an agreement to have the client's money managed with high risk in a Bank. The client denied having ticked the high-risk box, while the bank employee who managed the money of this account affirmed that this box was filled by the client. It was of interest for the mandating authority to determine whether the X-mark was affixed by the client or by the bank employee. Additionally, it was also asked whether the questioned X-mark was compatible with the writing of a right-handed writer or a left-handed writer.

We had at our disposal the original questioned document and some course of business reference documents from the client and the bank employee. We did not consider this reference material as adequate for comparison purposes and requested writing exemplars to be established in our laboratory. The client and the bank employee were called upon in turn and were requested to produce handwriting specimens during a single session, divided in three different tasks. The first task was to write, under dictation, a list of words and a text containing many lowercase letters *x* in different positions. In the second task the writers were asked to resolve a series of simple mathematical operations and to write answers in words, which included letters *x* (we expected that the writer attention would be focused on the resolution of mathematical operations rather than of the writing process, and this would be beneficial to capture writing habits). In the third task, the writers had to write X-marks in a series of boxes of different sizes printed on paper, to fill out grids of lottery tickets and a multiple-choice questionnaire. A number of 58 *x* was expected in task 1, 20 in task 2 and 76 in task 3, which represents a total of 154 *x* per writer. The actual number of *x* collected was slightly lower, due to spelling mistakes and boxes left blank. The session of each writer was video recorded to examine the construction of the letters *x* (i.e., for the determination of the stroke sequence) and if this construction is consistent through the session. For sake of simplicity, whatever the *x* is a letter or a cross in a box, it will be further designated as an X-mark.

We considered 8 theoretical possible constructions of an X-mark, depending on the stroke direction and sequence [Figure 1].

The questioned X-mark [Figure 2] was observed through magnification by using a stereomicroscope (Leica A60 magnification 5x to 30x) and a digital microscope (Keyence VHX-600, magnification 20x to 200x). We determined that the questioned X-mark was affixed with a blue ball point pen. The direction of the strokes of the questioned X-mark could be established, based on the striae left by the ballpoint (Ellen, 1997, p. 13) and inks deposits along the cellulose fibers following the methodology described in Devlin *et al.* (2015). The stroke sequence could not be determined on the questioned X-mark. According to these findings, the questioned X-mark corresponded to either construction #**3** or #**7**.

Examination of the video recordings revealed that the construction of X-marks was consistent among all tasks in a given writer. It proved that both the client and the bank employee used systematically and only construction #**5**. Both of them were left-handed writers.

Based on the information gathered on the construction of the X-mark, we could conclude that

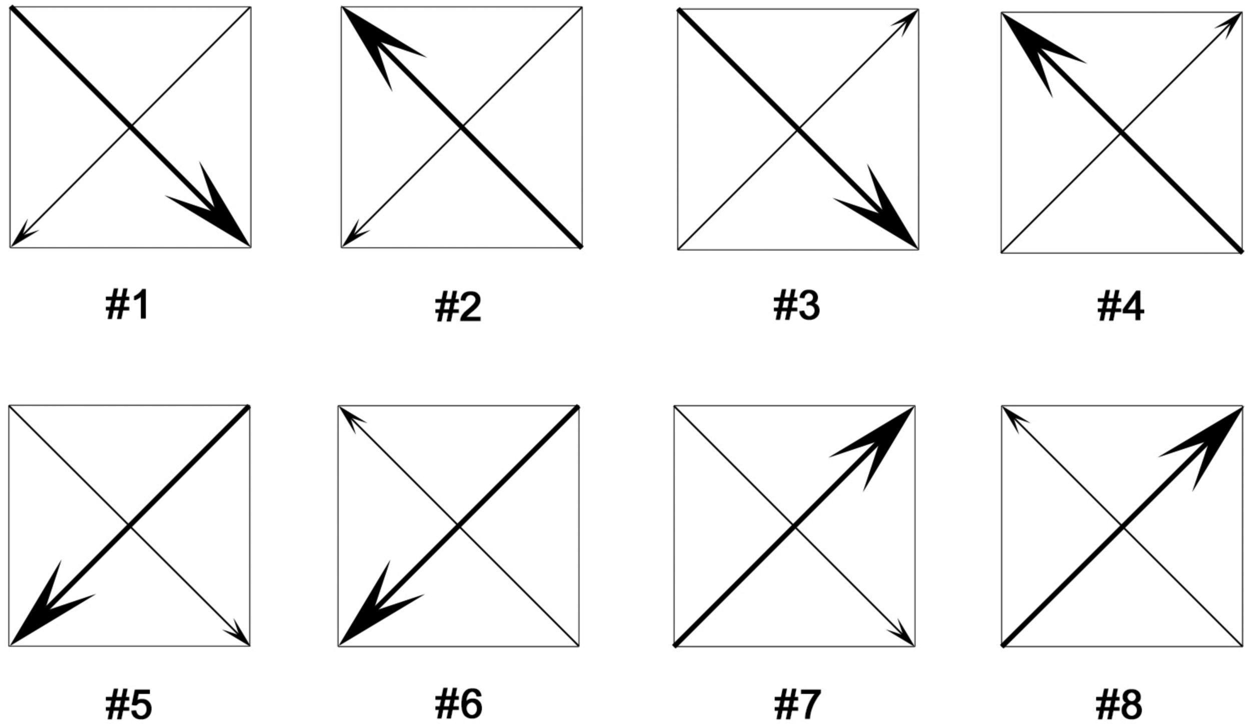

Figure 1. The 8 possible constructions of an X-mark. The points of the arrows show the stroke direction, while the size represents the stroke sequence: the heavy arrow being traced first, and the light one being traced second.

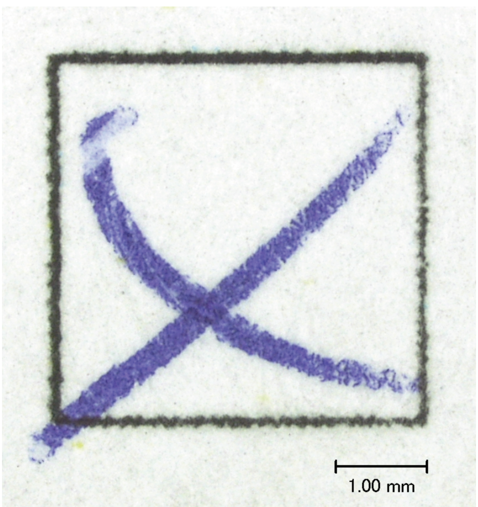

Figure 2. Illustration of the questioned X-mark.

both the client and the bank employee had a different letter construction than the questioned X-mark. These findings support that the questioned X-mark was neither written by the client, nor by the bank employee, but rather by some unknown person. It must be specified at this point that—based on the case information—it was assumed the questioned X-mark was naturally written: we did not consider the possibility that the writing was a simulation (of X-marks written by someone else) or a disguise (voluntary modification to escape identification). This assumption would be disclosed in our statement and, should this not be a reasonable assumption, then a new evaluation would be needed. We will see below how to quantify the strength of the evidence regarding the X-mark construction.

Empirical study

To determine whether the questioned X-mark was more compatible with a left-handed writer or a right-handed writer, an empirical study was carried out within a population of 100 writers of the School of Criminal Justice. Each subject was given a single sheet of paper with 20 printed boxes of size 4mm x 4mm and was requested to fill out the boxes with a blue ballpoint pen, while sitting at a table. The operator of the study was alone with the subject and could observe the sequence and the direction of the strokes. The construction (#**1** to #**8**) was identified and noted on the subject's sheet, together with handedness of the subject. The results of this study are reported in Table 1. It is worth noting that the construction adopted by each subject was consistent all along the 20 filled out boxes. In other words, each subject used only one construction for the X-mark.

Table 1. Results of the empirical study carried out within a population of 100 writers. #1 to #8 represent the 8 possible constructions of an X-mark, depending on the sequence and direction of the strokes (see Figure 1).

	#1	#2	#3	#4	#5	#6	#7	#8	Total
Left-handed writers	9	-	-	-	12	4	-	-	25
Right-handed writers	19	-	17	-	34	-	5	-	75
Total	28	-	17	-	46	4	5	-	100

As one can see in Table 1, constructions #**2**, #**4** and #**8** were not represented in our sample. The construction #**6** was only made by left-handed writers, while constructions #**3** and #**7** were only made by right-handed writers. Based on these findings, the questioned X-mark, of construction #**3** or #**7**, would be more expected in right-handed writers, than in left-handed writers.

Handedness evaluation

Due to indetermination of stroke sequence in our case, we use four construction classes: **A** (grouping #**1** and #**5**), **B** (grouping #**2** and #**6**), **C** (grouping #**3** and #**7**) and **D** (grouping #**4** and #**8**). Each class indeed groups a pair of constructions that only differ in their stroke direction. It is therefore relevant to consider only these four classes for the following developments (see Figure 3).

To help provide an investigative lead, we use the same metric suggested within the ENFSI[1] guideline on evaluative reporting (Willis *et al.*, 2015), and assign a likelihood ratio (LR) based on competing propositions. If the evidence (i.e., construction class) has value in this case, it will help discriminate the following propositions:

1. The questioned X-mark was affixed by a right-handed writer (R)
2. The questioned X-mark was affixed by a left-handed writer (L)

We will assign our LR based on the data acquired in our limited study and our knowledge. A LR is defined by the probability of the evidence if the first proposition is true, divided by the probability of the evidence if the second proposition is true. All probabilities being conditional, the numerator and denominator of our LR will depend on the information available to us. Before assigning values to these probabilities, let us think of the knowledge we have on distributions of constructions classes between right and left-handed writers. We consider that the data we have collected during our experiment comes to modify somehow our prior knowledge about the occurrence of given construction classes of X-mark in both populations of right and left-handed writers. We suppose that before carrying out the experimental study each construction class is equiprobable: that is, we have no reason to think that any of the construction class is more or less likely to occur than any other, whatever the population of right or left-handed writers. Therefore, we assign the same prior counts of 1 to each construction class. These prior counts can be viewed as a fictive set of initial observations that represents our prior knowledge on the distribution of each construction class (i.e. before the experimental

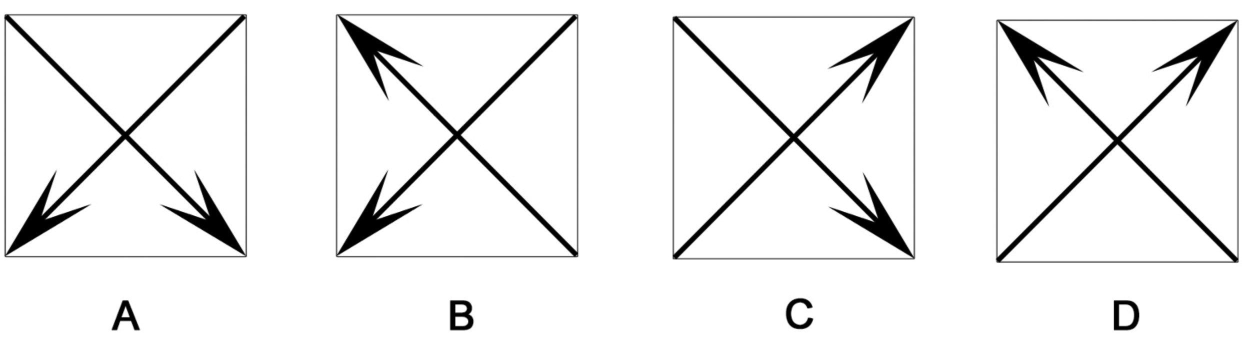

Figure 3. The four construction classes of an X-mark, based on the direction of the strokes. The points of the arrows show the stroke direction.

[1] European Network of Forensic Science Institutes

study). These prior counts, as well as the observed counts of the experimental study (i.e., our real observations) and the posterior counts are reported in Table 2. Posterior counts are the sum of the prior counts with the observed counts, i.e. the sum of the fictive initial observations and the real observations made during the experimental study. This procedure of updating distributions data was presented in Biedermann *et al.* (2009) and recently used in Samie *et al.* (2016).

Note that besides embracing a Bayesian approach for adjusting distributions, which is quite nice for a scientist who is engaged in this paradigm of uncertainty management, there is a practical advantage to consider the sum between prior and observed counts. It indeed resolves the problem of 0 observation data that make impossible the assignment of a LR. Indeed, construction class **C** was never observed in left-handed writers in our experimental study. The probability of observing such a construction if the writer were a left-handed writer would be 0, which could definitely not be an acceptable value for the denominator of our LR.

We are now ready to assign values to probabilities and evaluate the results. According to Table 3, the probability of observing an X-mark of class **C** if the writer is a right-handed writer, divided by the probability of observing an X-mark of class **C** if the writer is a left-handed writer, is:

LR = P(**C**|R)/ P(**C**|L) = (23/79)/(1/29) ~ 8.

This is based only on the data of our experimental study. One could argue that we should also add the two observations of our case (i.e. the observation from the client and the observation from the bank employee) as this is now part of our knowledge. Taking into account these observations, the LR would become: LR = (23/79)/(1/29+2) ~ 8. As these two observations belong to class **A**, their influence on the LR value is negligible. The situation would have been different if these observations were classified as **C**.

Our LR value means that our evidence (i.e., construction **C**) is in the order of 8 times more probable if the writer of the questioned X-mark is right-handed, than left-handed. Of course, this does not mean that it is 8 times more probable that the writer of the questioned X-mark is right-handed, rather than left-handed. As explained for example in Marquis *et al.* (2016), in order to obtain the posterior probability that the writer of the questioned X-mark was right or left-handed—i.e. given the evidence—we must combine our LR

Table 2. Prior counts, observations and posterior counts of the construction classes of X-marks in populations of right and left-handed writers. The observations result from the experimental study carried out for the needs of the case (see Table 1), grouping constructions #**1** and #**5** in class **A**, #**2** and #**6** in class **B**, #**3** and #**7** in class **C**, #**4** and #**8** in class **D**. Posterior counts represent the sum of the prior counts with the observations.

		Right-handed writers		Left-handed writers	
Outcome	**Prior counts**	**Observations**	**Posterior counts**	**Observations**	**Posterior counts**
A	1	53	54	21	22
B	1	0	1	4	5
C	1	22	23	0	1
D	1	0	1	0	1
Total	4	75	79	25	29

Table 3. Probabilities of observing a given construction class if the writer is a right or a left-handed writer.

Outcome	**Probability of a given construction class if the writer is right-handed**	**Probability of a given construction class if the writer is left-handed**
A	P(A|R) = 54/79	P(A|L) = 22/29
B	P(B|R) = 1/79	P(B|L) = 5/29
C	P(C|R) = 23/79	P(C|L) = 1/29
D	P(D|R) = 1/79	P(D|L) = 1/29

value with prior probabilities that the writer is right or left-handed. These prior probabilities represent the view of the trier of fact that the writer is a right-handed or a left-handed writer, given all the other relevant information of the case, without considering the forensic results.

Let us first consider the situation where the trier of fact does not favour any of the propositions, and considers it is equally probable that the writer of the questioned X-mark is a right or a left-handed writer. Given such prior probabilities of 50% (for the proposition of a right-handed writer) and 50% (for the proposition of a left-handed-writer), our LR of 8 moves the probability that the writer of the questioned X-mark is right-handed from 50% to 88.9%, and left-handed from 50% to 11.1%. The calculation of this example is given in Appendix. Details regarding the computation of posterior probabilities on the basis of prior probabilities and a LR value can for instance be found in Robertson *et al.* (2016) and Marquis *et al.* (2017).

Let us consider a more realistic situation where the trier of fact takes as prior probabilities the frequencies of right and left-handed writers in the general population, i.e. respectively about 90% and 10% (Huber and Headrick, p. 199). Given such prior probabilities of 90% (for the proposition of a right-handed writer) and 10% (for the proposition of a left-handed writer), our LR of 8 moves the probability that the writer of the questioned X-mark is right-handed from 90% to 98.6%, and left-handed from 10% to 1.4%.

Handedness misleading evidence

If a LR value is pointing towards a right-handed writer, what is the chance that the true writer was in fact a left-handed writer? It may be of interest to determine the proportion of misleading evidence, i.e. the proportion of cases where the proposition being supported by the data does not correspond to the true proposition. In such cases, the delivered information may lead the trier of fact on a wrong investigation way.

To address this issue, which is related to what is called the *robustness* of the LR (Taroni *et al.,* 2010, pp. 303-305), a series of LR values were derived from the data collected during our experimental study, based on the following procedure. Among the 75 observations of the right-handed writers, we have drawn each observation in turn as a questioned X-mark, while the other observations (i.e. the 74 remaining observations from the right-handed writers and the 25 from the left-handed writers) were taken as reference material. A LR value was generated for each draw. As we know that the questioned X-mark was taken from the right-handed writers, we should get a value supporting the proposition of a right-handed writer (i.e. LR >1).

The same procedure was applied within the 25 left-handed writers, taking as reference material the 75 observations from the right-handed writers and the remaining 24 from the left-handed writers. Again, as we know that the questioned X-mark was taken from the left-handed writers, we should get a value supporting the proposition of a left-handed writer (i.e. LR <1).

According to Table 2, the X-marks of right-handed writers included in our experimental study either belong to class **A** or **C**. If an X-mark of class **A** is taken from the right-handed writers population, the LR value[2] is $P(\mathbf{A}|R)/ P(\mathbf{A}|L) = (53/78)/(22/29) \sim 1$. If an X-mark of class **C** is taken from the right-handed writers population, the LR value is $P(\mathbf{C}|R)/ P(\mathbf{C}|L) = (22/78)/(1/29) \sim 8$. If the X-mark belongs to class **C**, the LR value correctly supports the proposition of a right-handed writer (LR>1). If it is of class **A**, then the results are neutral.

The X-marks of the left-handed writers included in our experimental study either belong to class **A** or **B**. If an X-mark of class **A** is taken from the left-handed writers population, the LR value is $P(\mathbf{A}|R)/ P(\mathbf{A}|L) = (54/79)/(21/28) \sim 1$. This value neither supports the proposition of a right-handed or a left-handed writer, which is expected as the proportion of this characteristic is about the same in both populations. If an X-mark of class **B** is taken from the left-handed writers population, the LR value is $P(\mathbf{B}|R)/ P(\mathbf{B}|L) = (1/79)\ (4/28) \sim 0.1$. This value correctly supports in the order of 10 times more the proposition of a left-handed writer (LR<1) rather than a right-handed writer.

Table 4 shows the counts for the four LR values in the right and left-handed writer populations.

The robustness of the LR provides an estimation on how many times a LR of a given magni-

[2] Note that the numerator of the LR is based on 53 posterior counts instead of 54, since one observation was drawn and taken as a questioned observation. Also, whether or not considering the observations made in our case on the two bank employees, our LR remains in the order of 1.

Table 4. Counts for the four observed LR values in both populations of right and left-handed writers.

LR value	Right-handed writers	Left-handed writers
~ 8	23	0
~ 1	54	22
~ 0.1	0	5

tude will point in the wrong direction (Taroni *et al.,* 2010, pp. 303-305). In our case we have obtained a LR value in the order of 8, which slightly supports the proposition that the questioned X-mark was written by a right-handed writer. According to Table 4, with the knowledge acquired here, with such a LR value, we never support the wrong proposition.

The results reported in Table 4 can also be used to estimate the proportion of misleading LRs, using the proportion of LR values lower than 1 and larger than 1. According to Table 4, no misleading result was obtained. Indeed, if the writer of the questioned X-mark was a right-handed writer, the correct proposition was supported—with a slight support (LR ~ 8)—in 30% of the cases (23/23+54), while a neutral result (LR ~ 1) was obtained in 70% of the cases (54/23+54). In the same way, no false positive result was obtained: if the writer of the questioned X-mark was a left-handed writer, the correct proposition was supported—with a slight support (LR ~ 0.1)—in 20% of the cases (5/22+5), while a neutral result (LR ~ 1) was obtained in 80% of the cases (22/22+5).

Most of the times, the LR value is neutral, which means that the evidence does not help discriminate between both propositions of interest. In a limited but non-negligible proportion of cases, the evidence proves to be valuable in that view. Nevertheless, our LR values are generally low, which means that the selected feature only provides a limited support towards the propositions of a right or a left-handed writer. In other words, this feature may generally be of limited relevance to discriminate between right and left-handed writers. This result is expected since a single feature only—based on the direction of two strokes—was considered.

Inference of writership with a matching hypothetical suspect

The LR value in the order of 8 pointing towards a right-handed-writer, together with suggestions of posterior probabilities of propositions, were communicated to the mandating authority. Let us assume that based on these results, investigations among right handed-writers of the bank were pursued. Another bank employee was suspected of having written the questioned X-mark. He was called upon to produce, under our request, reference X-marks following the same procedure as before. Let us further assume that this new bank employee uses only construction class **C**, i.e. the same as observed on the questioned X-mark. In this case, the evidence we would seek to assess is represented by the comparison results between the construction of the questioned X-mark and the reference X-marks written by the new bank employee. We may pose the following propositions to interpret the evidence, which respectively represent the views of the prosecution (H_p) and the defense (H_d):

H_p: The questioned X-mark was written by the new bank employee.

H_d: The questioned X-mark was written by an unknown person.

In this scenario, the LR is the probability of the evidence given that the questioned X-mark was written by the new bank employee, divided by the probability that it was written by another (unknown) writer. Note that we do not specify that the unknown person is a right-handed writer, although investigations were only carried out within this specific population to find a new suspect. First, it cannot be eliminated at first glance—by any observer—the possibility that the unknown writer is a left-handed writer. This means that the information pointing towards a right-handed writer is part of the expert's findings, which should not be included in the proposition, as thoroughly argued by Hicks *et al.* (2015). Second, the distribution of construction classes is not the same between right-handed and left-handed writers. It would be irrelevant and potentially misleading to include in the second proposition that the unknown writer is a right-handed writer. However, the trier of fact must understand that the fact that the suspect is a right-handed writer has been accounted for by the expert and should not be included in the prior beliefs for the propositions. This is necessary to prevent double counting this information.

If the questioned X-mark were written by the new bank employee, the probability of a correspondence of construction class would be very high (since we have observed that there is consistency regarding the class construction), say close to 1. If the questioned X-mark were written by another (unknown) writer, the probability of picking another writer by chance who uses a construction class **C** would depend on whether the writer is a left-handed writer (L, with a prior probability of 0.1) or a right-handed writer (R, with a prior probability of 0.9). Informed by the data reported in Table 3, the probability of the evidence if the proposition of the defense is true would be, in agreement with the rule of the extension of conversation (Lindley, 1991): $P(\mathbf{C}|H_d) = Pr(\mathbf{C}|L,H_d)*Pr(L|H_d) + Pr(\mathbf{C}|R,H_d)*Pr(R|H_d) = (1/29)*0.1 + (23/79)*0.9 \sim 0.27$. The LR is therefore 1/0.27, i.e. LR ~ 4. In other words, our evidence is in the order of 4 times more probable if the questioned X-mark was written by the new bank employee than if it was written by unknown writer.

The mandating authority is interested in the probability that the new bank employee wrote the questioned X-mark. This probability depends not only on the forensic results, but also on the other elements of the case. Based on the other information, the trier of fact has some belief that the new bank employee wrote the questioned X-mark: this is what is formally called a prior probability. This prior probability should not be based on the fact that the new bank employee is a right-handed writer, nor on the forensic examination provided, but only on the other elements of the case. Combining the LR value with prior probabilities of the propositions, we could inform the mandating authority that:

- If the trier of fact had good information to think that the suspect is the author, for example prior probabilities of $P(H_p)=0.9$ and $P(H_d)=0.1$, our LR value would move the probability that the questioned X-mark was written by the new bank employee from 0.9 (prior) to 0.98 (posterior), and the probability that the questioned X-mark was written by some unknown writer from 0.1 (prior) to 0.02 (posterior);
- If the trier of fact had no information to favor that the suspect is the author, then with equal prior probabilities of $P(H_p)=0.5$ and $P(H_d)=0.5$, the probability that the questioned X-mark was written by the new bank employee would move from 0.5 (prior) to 0.8 (posterior); and the probability that the questioned X-mark was written by an unknown writer would move from 0.5 (prior) to 0.2 (posterior);
- And, if the trier of fact had good information to think that the suspect is **not** the author, for example if prior probabilities of $P(H_p)$ were 0.1 and $P(H_d)$ 0.9, the probability that the questioned X-mark was written by the new bank employee would move from 0.1 (prior) to 0.4 (posterior) and the probability that the questioned X-mark was written by an unknown writer would move from 0.9 (prior) to about 0.6 (posterior).

Discussion

In this X-mark case, we were able to achieve a valuable conclusion based on stroke direction since this direction could be determined with certainty. While such a determination is generally possible if the writing instrument used is a ballpoint pen, it can be more difficult or even impossible with another type of writing instrument, for instance with roller pens or fiber tips. On another hand, while in our case we were unable to determine the stroke sequence, we have nevertheless considered worth reporting also the data that take into account this feature (see Table 1), based on our experiment. Such data can be highly beneficial for experts who might encounter a case where the stroke sequence can be determined. This sequence can sometimes be determined without ambiguity, especially in cases where the strokes of the X-mark are connected.

Huber and Headrick (1999, p. 149) pointed out the fact that the consistency of the X-mark execution is a key element to be considered. We suppose that a within-writer variation can be expected in various features (such as angle between strokes, position of the crossing points, etc.) but may not be likely to occur concerning the type of construction. In this respect, consistency was confirmed in all writers of the experimental study, as well as in both the client and the bank employee involved in the present case. This within-writer consistency could further be investigated through time or special conditions, which was not explored in this study.

Only one feature—the stroke direction—was considered and used in our evaluation (i.e., to assign our likelihood ratio). This is believed to be relevant to discriminate between right-handed writers and left-handed writers, as we assume this feature is one of the main, if not only, X-mark feature that may significantly differ between populations of right-handed and left-handed writers.

Nevertheless, the stroke direction proved to be of limited value to discriminate between right and left-handed writers. Improvement of handedness discrimination would be achieved by including features of other handwriting items (if present), such as direction of horizontal strokes of characters *f*, *t*, *A*, *E*, *F*, *H*, *T*, *Z*, *5*, *7* and direction in which circular structures are formed, especially in characters *o*, *O*, *Q* and *0* (Conrad, 2008).

The same feature, taken as its own, was then evaluated considering a different problem, thus a different set of propositions (i.e. to discriminate between a given writer—say Mr. A—vs another unknown writer). In such a case, the stroke direction is not the only relevant feature to be used. Other features, such as those described above (orientation of strokes, position of crossing between strokes, etc.) provide valuable information to discriminate between writers. The value derived only from the stroke direction data gives an order of magnitude of the LR for this single feature, but it represents a restricted view of the value of the observations made in X-marks. In casework, handwriting examiners may thus consider other features as well, since these additional features add value. This added value means that taking into account these additional features—and considering they do not differ between the questioned X-mark and the reference material from Mr. A—the LR value must logically be larger. This is true even if there are no specific survey on which to rely for these observed features. The examiners will thus combine quantitative data and traditional qualitative observations (using their calibrated knowledge, i.e. knowledge gathered in cases where the ground truth was known). An important point is that one should take into account the dependencies of the features assessed. When accounting for this, the more discriminant these features are between the writers, the larger the increase in the LR value. Here, in this setting devoted to the discrimination between writers, we would not expect a drastic increase of our LR (i.e., the increase would be assigned as less than a factor of two).

The experiment made during the case was carried out in two steps, which were not detailed in the present paper but will be briefly explained here. A first step included only 52 writers. This first dataset led to a LR value in the order of 3 for the discrimination between propositions of a right-handed vs a left-handed writer. At this point, we were wondering whether this small LR value was duly representative of the limited information contained in the stroke direction of X-marks, or whether this was due to our limited sample, which may not properly encapsulate the distribution of construction classes of X-marks. We have therefore made a series of simulations, leading to different distributions of construction classes, and different LR values. We have found that with a sample size of 100 writers, we could reasonably reach a LR value in the order of 10. Given such a potential gain and given the small cost (in time and money) of the task of collecting additional written X-marks, we have considered it was worth increasing the sample size. Increasing the sample size indeed appeared to be beneficial, and in agreement with our expectations, since a LR value in the order of 8 was obtained based on the full dataset of our experiment. We therefore consider it is a good idea, in cases where it can be disputed whether the sample size is sufficient, to make prognostics and investigate what could theoretically be obtained by increasing the sample size. If the impact of additional experiments on the value of the evidence is high enough compared to their cost, it can be worth increasing the sample size.

The procedure of evidence interpretation detailed in this paper may be followed by any examiner interested in applying a Bayesian approach on simple data collected for assessing the results of a given case, should this concern an X-mark or any other handwritten sign or letter. However, it should be underlined that the figures shown in this paper on the X-mark construction may not be representative of construction manners found in other countries. The distribution of construction classes may indeed depend on the country, especially because of different taught methods. Further experiments should be carried out to address this issue.

Conclusion

This paper shows that useful results can be obtained despite the apparent simplicity of the written entry to be examined. Of course, measurements or calculated frequencies are not mandatory to adopt a likelihood ratio approach for evidence interpretation, nevertheless this paper shows that following the ENFSI recommendations for evidence interpretation with numbers may require only a small dataset collected for the case needs. The LR as a metric appears convenient either in investigative or evaluative phases of the case file. The conclusion addressed on the handedness of the writer typically represents the investigative role of the expert, where the expert's conclusion can help narrowing down a population of putative suspects during the investigation phase. In the field of handwriting examination, the use of such a metric in investigative purposes—to distinguish between right-handed writers from left-handed writers—was already addressed in Taroni *et al.* (2011). On the other hand, the conclusion regarding the comparison of the X-mark with the handwriting of the bank employee represents the evaluative role of the expert and this is more commonly the phase where the LR is referred to (Willis *et al.*, 2015).

References

Biedermann, A.; Taroni, F.; Bozza S. (2009). Implementing statistical learning methods through Bayesian networks. Part 1: a guide to Bayesian parameter estimation using forensic science data. *Forensic Science International*, Vol. 193, pp. 63-71.

Cadola, L.; Margot, P.; Marquis, R. (2013). Are Simple Signatures So Easy to Simulate? *Journal of the American Society of Questioned Document Examiners,* Vol. 16, pp. 3-11.

Conrad, M. (2008). Left-hand writing vs. right hand writing. *Journal of the American Society of Questioned Document Examiners,* Vol. 11, pp. 19-27.

Devlin, A.; Graydon, C.; Strach, S. J. (2015). Line Direction Determination of Ballpoint Pen Ink Writing. *Journal of the American Society of Questioned Document Examiners*, Vol. 18, pp. 29-39.

Ellen, D. (1997). The Scientific Examination of Documents: Methods and Techniques. Second Edition. London: Taylor & Francis.

Foley, D. N. (1999). "X"/Cross marks—still used today. Presented at the annual meeting of the *Southeastern Association of Forensic Document Examiners*, Conyers, GA.

Found, B.; Rogers, D.; Rowe, V.; Dick, D. (1998). Statistical modeling of expert's perceptions of the ease of signature simulation. *Journal of Forensic Document Examination,* Vol. 11, pp. 73-99.

Hicks, T.; Biedermann, A.; De Koeijer, J. A.; Taroni, F.; Champod, C.; Evett, I. (2015). The importance of distinguishing information from evidence. *Science and Justice*, Vol. 55, pp. 520-525.

Hilton O. (1982). Scientific examination of questioned documents. Revised Edition. Amsterdam: Elsevier.

Huber, R. A.; Headrick, A. M. (1999). Handwriting Identification: Facts and Fundamentals. Boca Raton: CRC Press.

Lindley, D. V. (1991). Probability, in: Aitken, C. G. G.; Stoney, D.A. (Eds.), The Use of Statistics in Forensic Science. New York: Ellis Horwood, pp. 27–50.

Marquis, R.; Biedermann, A.; Cadola, L.; Champod, C.; Gueissaz, L.; Massonnet. G.; Mazzella, W.; Taroni, F.; Hicks, T. (2016). Discussion on how to implement a verbal scale in a forensic laboratory: Benefits, pitfalls and suggestions to avoid misunderstandings. *Science and Justice,* Vol. 56, pp. 364-370.

Marquis, R.; Cadola, L.; Mazzella, W.; Hicks T. (2017). What is the error margin of your signature analysis? *Forensic Science International*, Vol. 281, pp. e1-e8.

Osborn, A. S. (1932). Cases of interest—Cross-mark will set aside. *Law Notes*, Vol. 36.

Robertson, B.; Vignaux, G. A.; Berger, C. E. H. (2016). Interpreting evidence: Evaluating forensic science in the courtroom. Second Edition. Chichester: John Wiley & Sons.

Samie, L.; Hicks, T.; Castella, V.; Taroni, F. (2016). Stabbing simulations and DNA transfer. *Forensic Science International: Genetics,* Vol. 22, pp. 73-80.

Taroni, F.; Bozza, S.; Biedermann, A.; Garbolino , P.; Aitken, C. (2010). Data analysis in forensic science: A Bayesian decision perspective. Chichester: John Wiley & Sons.

Taroni, F.; Marquis, R.; Schmittbuhl, M.; Biedermann, A.; Thiéry, A.; Bozza, S. (2011). Bayes factor for investigative assessment of selected handwriting features. *Forensic Science International*, Vol. 242, pp. 266-273.

Welch, T. (1999). The mark and obliteration of the "X". *International Journal of Forensic Document Examiners*, Vol. 5, pp. 245-248.

Willis, S.; Mc Kenna, L.; Mc Dermott, S.; O' Donnell, G.; Barrett, A.; Rasmusson, B.; Höglund, T.; Nordgaard, A.; Berger, C.; Sjerps, M.; Molina, J. J. L.; Zadora, G.; Aitken, C.; Lovelock, T.; Lunt, L.; Champod, C.; Biedermann, A.; Hicks, T.; Taroni, F. (2015). ENFSI Guideline for evaluative reporting in forensic science. http://www.enfsi.eu/news/enfsi-guideline-evaluative-reporting-forensic-science

Appendix

According to the Bayes' theorem, the ratio between probabilities of the propositions (i.e. odds) of a right-handed writer (R) and a left-handed writer (L) given the evidence (C), i.e. the ratio of posterior probabilities of the propositions, is given by the ratio between prior probabilities of the propositions multiplied by the value of the likelihood ratio. As probabilities are conditional on what we know, what we are told and what we assume, we indicate this by the information *I*.

Ratio of posterior probabilities of the propositions = Ratio of prior probabilities of the propositions × LR

$$\frac{P(R|C,I)}{P(L|C,I)} = \frac{P(R|I)}{P(L|I)} \times \frac{P(C|R,I)}{P(C|L,I)}$$

Taking the example of equal prior probabilities of 0.5 for P(R) and 0.5 for P(L), our prior odds will be 1:1. With our LR value of 8, the value of the ratio of posterior probabilities of the propositions is:

$$\frac{P(R|C,I)}{P(L|C,I)} = \frac{0.5}{0.5} \times 8$$

$$\frac{P(R|C,I)}{P(L|C,I)} = \frac{8}{1}$$

We consider that people are either left-handed or right-handed (so we discard ambidexter persons). These are the only two possibilities considered. We know then that the sum of these probabilities equals 1:

$$P(R|C,I) + P(L|C,I) = 1$$

Using substitution of the latter in the former, we obtain:

$$\frac{P(R|C,I)}{1 - P(R|C,I)} = 8$$

Therefore:

$$P(R|C,I) = \frac{8}{9}$$

$$\Pr(R|C,I) \cong 0.9$$

And:

$$P(L|C,I) = 1 - \frac{8}{9}$$

$$P(L|C,I) \cong 0.1$$

We can also use the known relationship between the odds and the probability, as shown in Marquis *et al.* (2017): if the odds are *a to b*, then the probability can be calculated using the following formula: $Pr = a/(a + b)$. In our case, the odds (i.e. the ratio of posterior probabilities) are 8 *to* 1 (in favor of the proposition of a right-handed writer), so the posterior probability of a right-handed writer is: $P(R|C,I) = 8/(8 + 1)$.