Journal of the American Society of Questioned Document Examiners, Inc.
Volume 23, Number 1, June 2020

An Evaluation of Cursive and Hand Printing Class Characteristic Significance in Limited Collected Writing Samples

Douglas A. Ridolfi, M.S.

Coordinator of Forensic Chemistry, Buffalo State College, Chemistry Department, 1300 Elmwood Avenue, Buffalo, NY 14222.

The paper describes the results of an exercise that trainees can use to develop their own criteria for distinguishing among handwriting features, apply simple statistical methods to correlate features and develop a scheme for describing these characteristics that allow students to rate the degree of agreement between similar writings. By using a classification system involving letter shape, slant, height ratio, etc., one can subdivide the group of writings by roughly half with each feature to arrive at the closest match to a given model.

Key words: Chi square, correlation, handwriting and hand printing classification.

Introduction

There has been an ongoing need for many years to utilize statistical methods in handwriting comparisons. The primary difficulty with the application of statistical methods is that writing is a dynamic, changing process. Classification protocols and statistical methods use discrete characteristics (primarily letter shapes). A wealth of information that is readily observable but continuous over a wide range (pen pressure variation, line quality, etc.) often goes unevaluated (Hilton, 1958). To this end, researchers have deconvoluted the process by which a questioned document examiner examines and compares questioned and standard writings that include (a) extraction of common letter forms and features, (b) having a method or nomenclature to describe these letter forms, (c) have questioned document examiners tabulate these forms in either lists of software data bases, (d) determine the frequency of occurrence of these features, use statistical models to determine frequency or correlation of features, and (f) use these features and statistical frequencies to infer the individualizing potential of these characteristics when arriving at an opinion of identity or non-identity or opinions in between (Srihari, 2013).

This project was an attempt to evaluate the quantity of information that can be captured by comparing letter forms and other discrete features in a limited collected writing sample. Simple statistical methods were used to evaluate the independence of selected writing characteristics. This study will hopefully serve as an extension of some of the early works of Harris (1958) on how much people write alike and the Muhlberger, et al. statistical study of handwriting characteristics (Muhlberger, et al, 1977) and point out the importance of correlation of features to each other in addition to the tabulation of feature occurrence (Johnson, et al., 2017) intended for students and trainees who desire some experience in simple statistical methods for evaluating handwriting.

Handwriting Development and Legibility

Before any classification or characterization of handwriting can begin, it would be appropriate to briefly look at the origin of our Roman alphabet and some common aspects of letter formation. Methods of handwriting instruction will also be briefly examined.

There are three theoretical considerations in the design of most modern alphabets. The first is that it covers all of the phonemes (i.e., the minimum acoustic constituents) in the language; the letter shapes should be restricted to a total of between 20 and 30 symbols for easy recall; the individual letter shapes must not be required to represent several sounds, that is, their acoustic identities must be fixed and unchanging. The Greek alphabet, later adopted by the Romans, was the first alphabet to satisfy all these conditions (Havelock, 1976). Letter combinations can form graphemes which are representative of sounds and are considered the smallest units of writing that correspond to a sound such as "sh" in the word "shake." Some identification software such as Sciometrics, LLC *Flash ID* software, uses algorithms to compare the shape and overall topology of these letter combinations in order to compare similar writings.

Initially, the Roman alphabet had 21 letters. After 600 years of changes and additions due to differences in pronunciation between the two languages, the official Roman alphabet was established. When the Romans conquered Greece in the first century B.C., this 21-letter alphabet received two more Greek symbols (Y and Z). The Medieval addition of three more letters (U, W, and J), which were differentiations of existing Latin letters, brought the total to the current 26 (Rhodes, 1978).

Around the time of the Middle Ages, the transition from manuscript writing to cursive writing took place. Cursive writing gained in popularity because of its more rapid execution and it soon became the dominant writing form among literate societies.

Throughout the years, various writing styles have been introduced into American schools.

Until around 1920, only cursive writing was taught in this country. Cursive writing is characterized by running or flowing lines with strokes joined within the word and angles rounded. Manuscript writing, introduced around 1920, is writing that utilizes modified forms of the printed letter, simplified and without ornamentation with the letters in a word not connected by strokes. This is commonly referred to as "printing" (Templin, 1958). Today, manuscript writing is taught in grades 1 and 2 and cursive is introduced in grade 3.

Kindergarten is a time when the child's basic motor skills are developed by involving the child in exercises such as drawing circles, horizontal, vertical, and diagonal straight lines (Lawton, et al., 1980). As Purtell summarized in his 1980 article:

> All the methods of instruction for manuscript writing follow the same procedure. Four basic strokes are used. The horizontal stroke is always made from left to right except in the return stroke after forming the curved arc at the bottom of the letters D, P, B and the middle of the letter R. Vertical strokes are made from the top down… Also, slanted lines, such as the A, V, and W are made from top to bottom. All circles are made counterclockwise except for the bowls of the letters b and p. Other letters containing clockwise arcs are h, j, m, n, r and the bottoms of the S. Regarding size and proportion, the lower case is a is considered one unit high. However, the capital letters, ascender and descender portions of the letters will be 75 to 80 percent larger, except for the letters d and t which are only 50 percent taller. These general observations hold fairly true of all systems studied.

The existence of similar writing styles coupled with penmanship instruction in the elementary grades has created broad similarities in the writings of most people educated in this country. In the elementary grades, correct formation of letters, connecting strokes, letter spacing, and overall legibility are emphasized rather than speed. Grades 4 and 5 are a time to further refine cursive script.

An extensive study of the handwriting ability of younger children was performed by Amy Tan in 1981. She found that children aged 3.3 or younger could only scribble while some children 3.6 or older could print a few letters and numerals. Most of the four-year old children could print their first names and print or copy most numerals and letters requested. Most could copy simple words using disconnected strokes. As the children got older, they made more continuous strokes, fewer discontinuous strokes and their writing decreased in size. The overall developmental pattern of children's handwriting behavior from age three to six involved the following sequence stages: 1) controlled scribbles; 2) discrete lines, dots, symbols; 3) straight line and circle upper-case letters; 4) upper case letters; 5) lower case letters, numerals and words.

Difficulties in writing legibility often become noticeable at the junior high school and lower grade levels. The student now relies on writing as a means of expressing thought rather than an exercise in and of itself. As a result, more emphasis is placed on speed and less on good letter form. Increased speed often results in cutting corners and the correct formation of letters suffers.

One of the first major studies of writing legibility was by Freeman (1918). In this study of writing movements, he divided writing into three levels of complexity:

1. Simple extension and flexion of one finger.
2. Straight pen movement using several fingers.
3. Complex combinations of finger, wrist, and arm movements.

In general, horizontal movement is wrist motion as is diagonal movement from lower left to upper right and back. Vertical movement is primarily a finger movement with occasional assistance from the in and out movements of the arm. Top to bottom movement is primarily fore finger and bottom to top is primarily middle finger. In rapid writing, the wrist follows the path of least resistance and tends to slur letters at the ends of words. Complex details in small letters become less distinct especially if the proper formation involves rapid changes in speed and pen direction. Fast writers tend to simplify letter forms. Angular, complex forms which require pauses, air movements or sudden angle changes tend to be most deteriorated.

These modifications in writing speed have resulted in legibility problems. A person writing in cursive form can write an average of 50% faster than someone writing in manuscript, which gives an average speed of about four letters per second, cursive (Wing, 1979). This increase in speed is the chief reason for illegible writing.

Newland, in his 1932 study found that four types of difficulties in formation of letters caused over one-half of illegible writings: 1) failure to close letters; 2) closing looped strokes; 3) looping non-looped strokes; 4) using straight-up stokes rather than rounded strokes. Writing "e" like an "I" resulted in 15% of all illegible writings and malformation of "I" accounted for at least 55% of capital letter illegible letter forms. Twelve percent of illegible letters involved the letter "r." the letters a, e, r and t together contributed to about 45% of illegible writing. Adults were found to write three times more illegible than elementary school children.

The net result of "cutting corners" is to deviate from standard copy book form which leads to greater individuality and variance in handwriting. Variations among different writers has led to greater ability to discriminate between people but has complicated attempts to develop workable handwriting classification systems.

Principles of Handwriting Identification

The Daubert decision involving handwriting comparison (*U.S. v. Starzecpyzel*) has brought some of the principles of handwriting comparison under critical scrutiny. The two fundamental principles brought out at the hearing were: 1) no two people write the same way (inter-writer differences) and 2) no person will write exactly the same when repeating the same segment of writing (natural variation).

The first principle is difficult to evaluate unless one has examined many writers who have written comparable material. The study by Harris (1958) was one of the first attempts at such an evaluation. He studied signatures from the Los Angeles County voters Registration and noted how closely people with the same last name (Smith) wrote their signature. He found that what one may regard as unusual letter forms appeared in the population sample more than one might expect, if the sample is sufficiently large. When comparing these unusual letter forms in two writers, their method of construction must agree as with any other letter. Slight but persistent differences in slant, size, shading, letter forms and movement between two specimens of handwriting is a strong indication that they are by different writers.

The importance of line quality should be emphasized. Line quality can be judged by the regularity of the line of writing, smoothness of curved letters and smoothness of transition from one stroke to the next (Harrison, 1966). People who write alike have a great natural potential to imitate each other's writing.

When writing between two people is compared, we are relying on an evaluation of similarities and dissimilarities. A questioned document examiner notes the stroke sequence and how the strokes are used to connect and form letters. The examiner will note spacing, size, slant, shading, speed, shape, and slope. These factors combined provide a clue to the skill level of the writer which is a subjective measure of the artistic sense of the writer and a measure of the level of variation in the writing and the amount of pen control exerted. The style of the writer is a global mixture of traits that may involve the use of capital letters, printing, cursive writing, a mixture of printed and cursive letters, types of connecting strokes and the manner in which letters are connected (Kam, 1994).

The questioned document examiner notes these characteristics and makes a mental note of total similarities and dissimilarities. Differences can be the deciding factor despite extensive similarities, especially if the differences are repeated and basic in nature. Similarities, sometimes even a large number, are to be expected since most people educated in this country were taught similar copy book form.

Hilton noted in 1983, "…writing habits fall into two general and somewhat overlapping groups—class and individual characteristics. Class characteristics…are those common to a number of writers and may result from such influences as the writing system studied…Individual characteristics are more or less peculiar to a specific writer…Without true individuality in personal writing habits we must assemble a combination of distinctive habits." This recognition of small, distinctive and individual traits and their combination is what is relied on by most experienced questioned document examiners and not on the overall pictorial appearance, as is common with lay observers.

Brit and Mensh in their 1943 study, noted that roughly 60% of a group of 181 college students made some initial error in identification of their own handwriting. The factors that they relied on were (in order): 1) general appearance; 2) appearance of certain letters and numerals; 3) pressure of writing; 4) slant of writing; 5) size of writing; 6) method of writing (Palmer, Spencer, etc.)

Regardless of what factors are relied on, it is the discrete factor such as letter shape that is easiest to capture for purposes of classification. Classifications applied to writing may have value in sorting through numerous submitted writings to identify a writer and have value in the statistical evaluation and correlation of features for further study. Application of systematic classification can eliminate some of the uncertainties of verbal description and help ensure that classifications and numerical data can be verified by other observers (Baxter, 1973).

Classification Systems

One of the first attempts to identify two writers by statistical evaluation of similarities and dissimilarities was a study of the Junius letters by Charles Chabot. He reported an exhaustive study of characteristics in a series of anonymous letters and recognized that "It is in their cumulative effect that similarities of a common character have weight." He noted, for example, 13 variants of the lower case "r" and such features as:

1. The coincidences…of numerous and diversified formations of the letter r and of the circumstances under which particular formations only were employed…
2. The various modes of terminating words.
3. The mode of inserting miswritten and omitted letters.
4. The datings and openings of letters.

Each gave a preference to r with its shoulder to the left in all cases where it follows either of the vowels a, e, I, o, or u with a strongly marked preference after the vowel a. There are 487 exceptions to this general rule in 1646 instances in Junius and 483 exceptions in 1790 instances in Francis [chief suspect]. Of these exceptions, 250 in Junius and 225 in Francis occur in the final letters of words (Chabot, 1871).

This was typical of the level of detail in the observation of writing features reported by Chabot. He recognized basic elements of disguise in the first series of writings and was able to tie these early writings in with subsequent, less disguised writings by the carryover of dominant characteristics that were still evident.

Although not strictly relying on statistical frequencies, there have been several attempts over the years to develop handwriting classification systems designed for the purpose of forgery and bad check investigation. All have met with only limited success because of natural variation in subjective evaluation of the same characteristics between one observer and another.

One of the first such attempts at classification was the method proposed by Lee and Abey (1922). They proposed a classification of cursive handwriting based on:

1. Form (angular, rounded, eyed).
2. Skill.
3. Connections (capitals connected, capitals disconnected, small letters connected).
4. Shading.
5. Movement (finger, compound, forearm).
6. Embellishments.
7. Terminals (upward, horizontal, downward).
8. Slant (less than 60 degrees, 60 to 80 degrees, more than 80 degrees).
9. Width of small letters.
10. Spacing of small letters within words.

11. Speed as determined by the quality of the stroke.
12. Proportion of single-space letters to capitals (less than one-third, one-third, more than one-third)

Using three classes within each of the major classes gave 9, 840 divisions. This was considered adequate for a collection of 100,000 checks, if equally divided between classes.

The difficulty in using such a classification system becomes evident when more than one reader/classifier attempts to input entries. The study by Eldridge, et al., (1984) on variability of selected features in cursive handwriting found an average disagreement rate of about 17% in classification of handwriting characteristics between different classifiers and a 7% disagreement rate when the same classifier went back and reclassified his previous classifications.

Eldridge's study looked at the six letters, d, f, h, k, p, and t. He found that on any given letter and feature combination, between 15 and 40% of the writers employed no fixed pattern. In such cases, a writer's letter formation was not consistent enough to be classified in any single category for that feature.

Hardcastle, et al. (1986) attempted to develop a computer-based handwriting classification system. The first part of their classification system relied on subjective, general features such as slope, connections, angularity, etc. The second part incorporated extensive subdivisions of letter forms in order to maximize the individuality of the category that the feature was assigned to. As with similar classifications, same-writer variability was one of the most difficult problems to contend with.

There are, however, some aspects of writing that tend to be less variable than letter form. Wing and Nimmo-Smith (1987) studied the height ratio of the letter combination "el" is such words as "bell", "well," etc. Even though height ratios are based on an underlying continuum rather than a discrete value such as letter form, the height ratio of two letters often show similar ratios even when a writer varies the absolute letter height, width, spacing or slope.

These studies emphasize that more practical classification systems are derived from using categories that provide the maximum distinction between types rather than the maximum number of categories for each letter. Systems that rely on features which are continuously variable over a wide range such as the degree of letter curvature or absolute letter height measurements should be avoided.

There have been fewer attempts to classify manuscript writing. Among the more noteworthy has been the method proposed by Livingston (1959). He based his classification on:

1. small letter or lower-case printed forms in half or more printed letters not counting the first letter of each capitalized word.
2. Capital letter or upper case in more than half of the letters.
3. Slant less than and slant greater than 65 degrees in at least three-fourths of the letters.
4. Formation of capital *A* (angular top, flat, or rounded top).
5. Capital *E* (long or short midstroke).
6. Capital *M* and *W* (high or low centers).
7. Variations of capital *R*, *N*, and lower-case *r*, *n*, *t* and *a*.

Springfield Study

The field of questioned documents has long been criticized by the courts and scientists in other disciplines for its lack of statistical evaluation of writing characteristics. There has been little critical evaluation of the frequency of given characteristics, the independence of writing feature combinations or the number or type of characteristics needed for an identification.

This study was conducted in an attempt to determine the frequency of easily classifiable characteristics; if there are any features that can divide handwriting into categories; if similar criteria for classification can be applied to both cursive and manuscript writing; and if there is a correlation between characteristics studied.

The sample consisted of handwritten and hand printed letters addressed to the State Archives in Springfield, IL. The letters were received between October 1985 and October 1986.

Three hundred eighty hand printed envelopes from 43 states and 1130 envelopes in cursive writing from 49 states were selected. All duplicate letters were culled out before the final group was selected. In the hand printed series of 380, one writer wrote 18 times. His letters were used as a model on which to develop a classification that would pick out other writers who wrote more closely to him. The 1130 cursive written envelopes were compared to one model which had been selected from a writer who had written eleven times.

With the cursive writing samples, the word "Springfield" was chosen as the model word for comparison (Figure 1). It is the longest single word in the address and has ten different letters. The words "Springfield" and "Archives" were selected for the hand printed samples because of

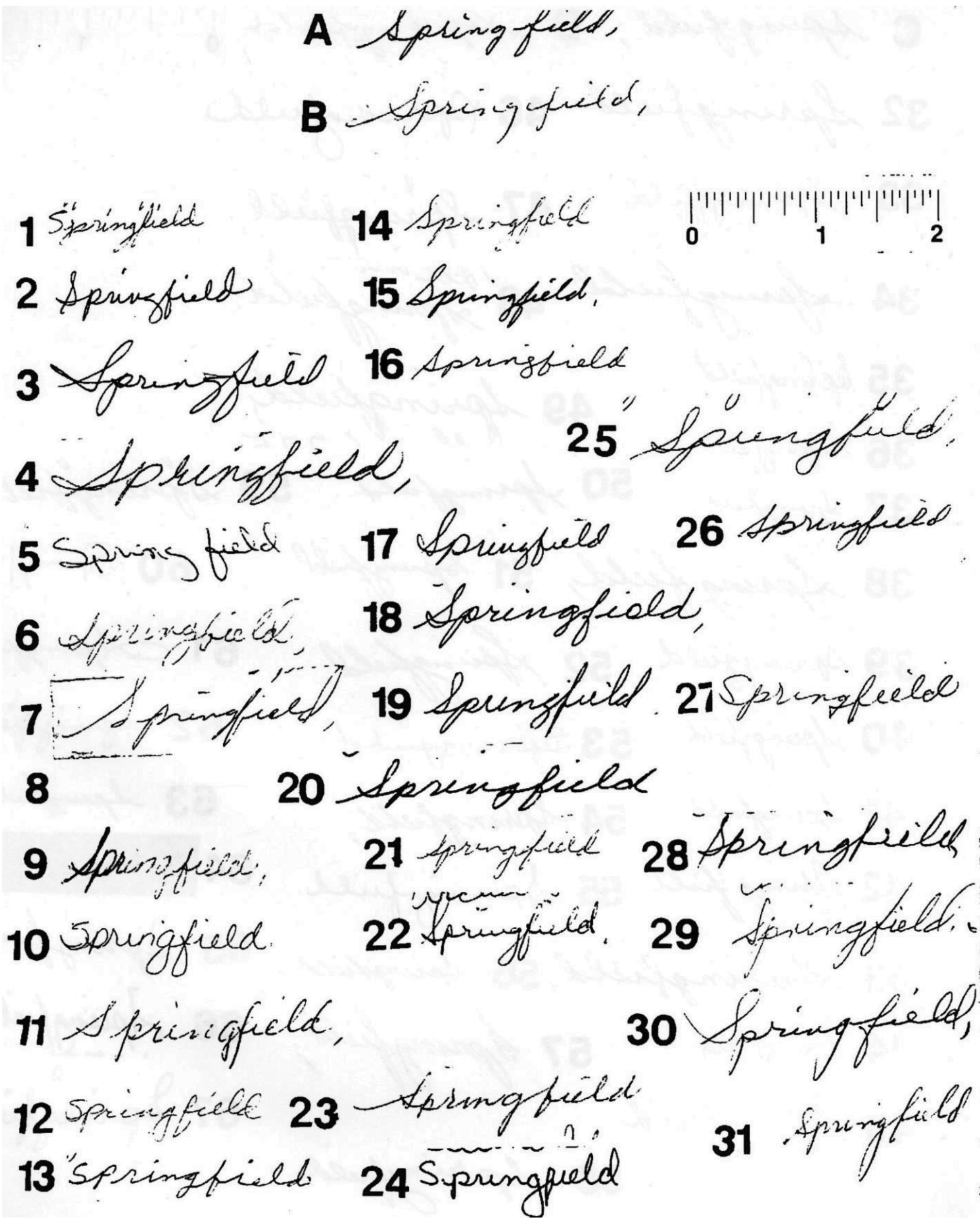


Figure 1. Cursive samples of "Springfield"

the anticipated fewer characteristics in common with the model printed in all block capitals (Figure 2). Classification was problematic. Many of the block capitals were not directly comparable to the lower-case letters that most other writers used within the two words.

The average English word length is 4.3 letters. About two percent of English words have 11 or more letters. The 14 letters in "Springfield Archives" accounts for about 70% of the letters of the alphabet in common use (Hall and Hardcastle, n.d.).

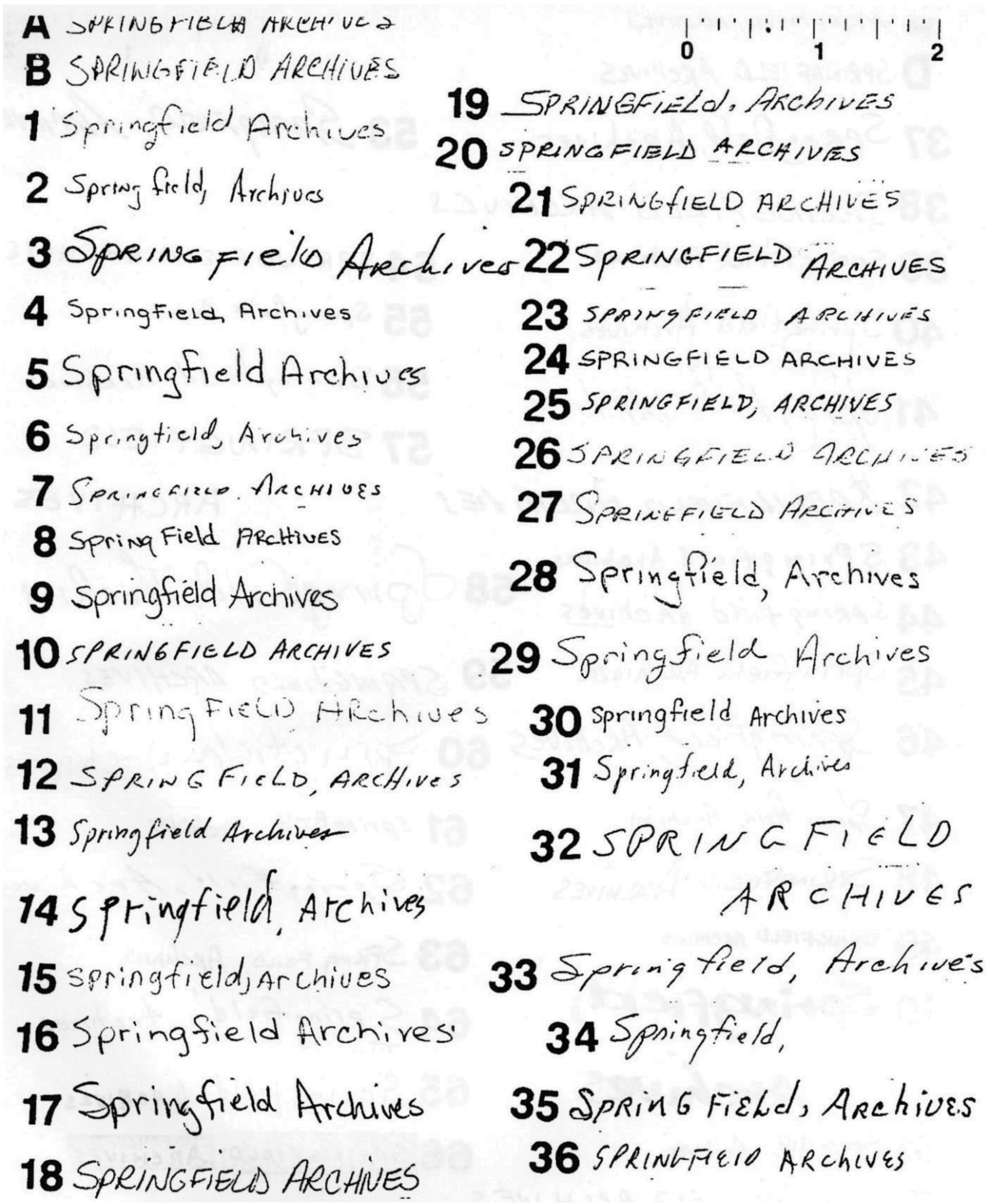


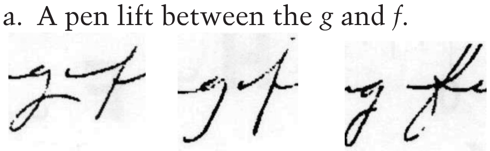
Figure 2. Handprinted samples of "SPRINGFIELD ARCHIVES"

Once all duplicate envelopes had been separated out (other than the duplicates in the models), the individual samples of "Springfield" were cut from each envelope and pasted onto a paperboard mount along with an identifying number corresponding to the envelope from which it was cut. The duplicates in the model were given a letter designation. The individual entries were arranged alphabetically by state from Alabama to Wyoming and alphabetically by city within each state as determined by the return address, simply for ease of classification.

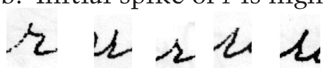
I looked at the eleven models of "Springfield" written by one person and developed a set of criteria that described a feature common to at least most of the model writings. The following twelve characteristics and their letter designations were used for classification (Figure 3):

There were no features of the model that could be considered rare or one of a kind. I concentrated on discrete values such as one letter characteristic being longer or higher in relation to another rather than on absolute height. I also avoided changes over a continuum such as letter curvature or anything that required measurements such as absolute values in slant. Ideally, features should be chosen to give as close to a 50:50 split as possible with half of the samples displaying the chosen feature and half that do not. Each of the 12 characteristics were counted among the 1130 cursive writing samples and their features tabulated. Results are shown in Table 1:

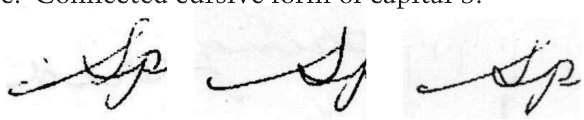
a. A pen lift between the *g* and *f*.



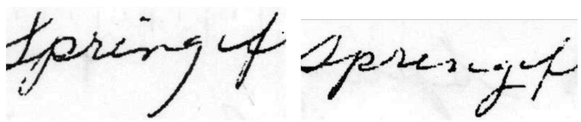
b. Initial spike of *r* is higher on left shoulder.



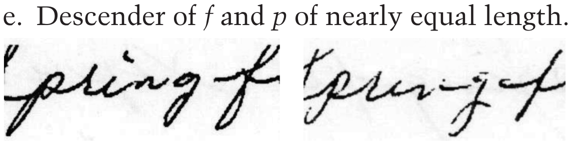
c. Connected cursive form of capital *S*.



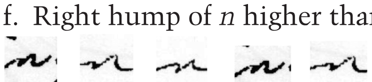
d. Loop of g longer than descender of either *f* or *p*.



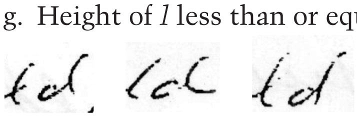
e. Descender of *f* and *p* of nearly equal length.



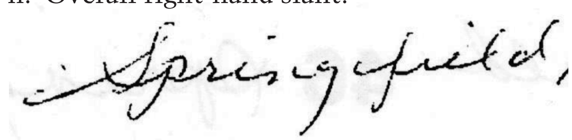
f. Right hump of *n* higher than left hump.



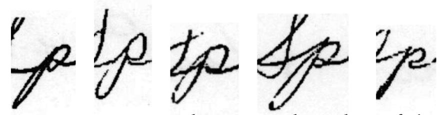
g. Height of *l* less than or equal to height of *d*.



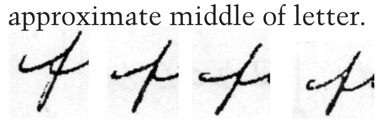
h. Overall right-hand slant.



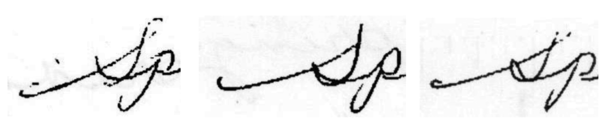
i. Open loop present in descender of *p.*



j. Connecting knot and spike of *f* in approximate middle of letter.



k. Approach stroke of *S* equal in length to at least ½ of height of letter and curved upward.



l. Apex of *r* equal to or higher than top of *p*.

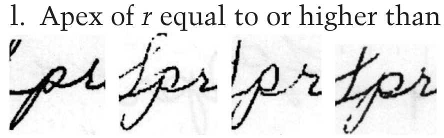


Figure 3. Images of chosen criteria.

Table 1. Percentage of handwriting characteristics.

Feature	Number of Times Observed	Percentage of Samples With Feature
A	322	28.5
B	476	42.1
C	765	67.7
D	452	40.0
E	457	40.4
F	452	40.0
G	932	85.5
H	904	80.0
I	785	69.5
J	397	35.1
K	234	20.7
L	878	77.7
Ave.	588	52.0

Table 2 was constructed listing each of the 1130 entries. Each feature found in an item in common with the model "Springfield" was listed. The following results were obtained:

Table 2. Frequency of observations of common features

Features in Common	Number of Times Observed
12	0
11	5
10	26
9	60
8	156
7	255
6	231
5	189
4	105
3	58
2	21
1	2
0	0

These results when graphed gave roughly a bell-shaped curve with seven characteristics, the number of characteristics in common with the greatest number of writings. This normal distribution permits the use of standard statistical treatment to evaluate the data.

Writings were next sorted by visual inspection and the number of features in common with the models were counted. The average number was eight characteristics in common. Entry number 183 had eleven characteristics in common with the model and bears a good resemblance on visual inspection. Entry 32 shows the greatest pictorial similarity when picked out visually. It has ten characteristics in common with the model writing. These are illustrated in Figure 4:

Entries were next sorted by characteristics. A batch of writings were sorted by looking at all writings starting with the most frequently appearing characteristic. In the sample of cursive writing, it was the height of the *l* less than or equal to the height of the *d* which was present in 85.2% of the writing. Samples sorted by this criterion were next divided into smaller groups by using the next most frequent criteria, overall right-hand slant, appearing in 80.0% of the samples. This led to the following subdivisions with letter designations from the features in common and number of times observed as seen in Table 3:

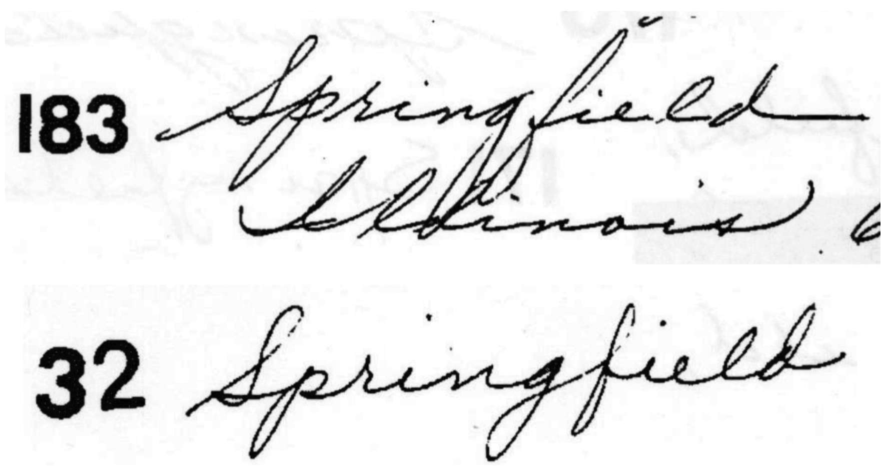


Figure 4. Writings sorted by visual inspection and the number of features in common.

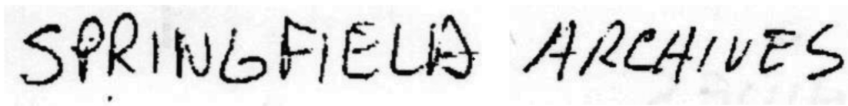
Table 3. Results of sorting criteria.

Sorting Criteria	Probability of Combination Occurrence Calculation	Number of Times Combination Observed	Frequency of Combination Occurrence
G	$(½)^1 = 0.5$	932	1/1.2 (0.833)
GH	$(½)^2 = 0.25$	739	1/1.5 (0.666)
GHL	$(½)^3 = 0.125$	563	1/2 (0.500)
GHLI	$(½)^4 = 0.0625$	404	1/2.8 (0.357)
GHLIC	$(½)^5 = 0.0312$	303	1/3.7 (0.270)
GHLICB	$(½)^6 = 0.0156$	161	1/7 (0.143)
GHLICBE	$(½)^7 = 0.0078$	68	1/16.6 (0.060)
GHLICBED	$(½)^8 = 0.0039$	35	1/32.3 (0.031)
GHLICBEDF	$(½)^9 = 0.0019$	17	1/66.5 (0.015)
GHLICBEDFJ	$(½)^{10} = 0.0009$	8	1/141 (0.007)
GHLICBEDFJA	$(½)^{11} = 0.0005$	1	1/1130 (0.001)

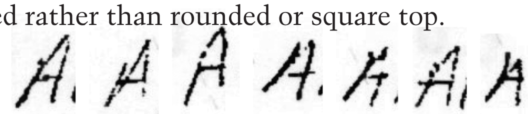
I used characteristics which I hoped would give a 50:50 distribution in the sample. Protocol was developed on the assumption that in the absence of an actual count, each characteristic ideally has a 50% chance of being present. In fact, the average division as one adds a characteristic is to eliminate 44% of the entries as each new feature is added. If each factor has a probability of ½ of occurring, then the two factors, that are independent of each other, have a (½) x (½) = ¼ chance of occurring. The probability of ten factors appearing together in one sample is then $(½)^{10}$ or one chance in 1024 of occurring.

Data for the 380 printed entries was captured in much the same way. I used the words "Springfield" and "Archives" because I anticipated that the printing might not be as rich in characteristics as cursive. The following were selected as illustrated in Figure 5:

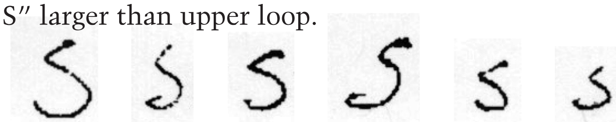
a. All capital letters or mix of capital letters greater than ½ all letters.



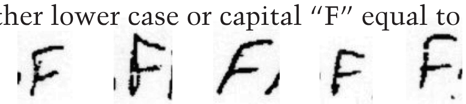
b. Capital "A" pointed rather than rounded or square top.



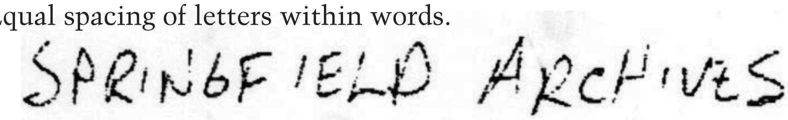
c. Lower loop of "S" larger than upper loop.



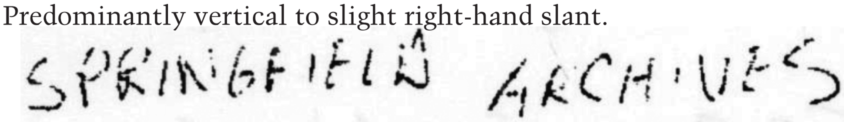
d. Lower cross bar or either lower case or capital "F" equal to or shorter than top cross bar.



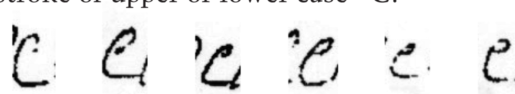
e. Equal spacing of letters within words.



f. Predominantly vertical to slight right-hand slant.



g. Recurve in upper stroke of upper or lower case "C."



h. Initial vertical stroke to capital "A" at least partially separate from main body.

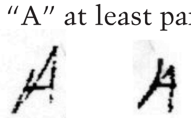


Figure 5. Images showing handprinting characteristics.

There were no features in the model that could be considered rare or one of a kind. All letters in the model were block capital letters which made direct comparison to lower case letters in the individual writings difficult and limited identification features that could be used. Discrete features such as letter shape and stroke formation were noted. A count of the eight features among the 380 samples were tabulated and their frequencies calculated. The results are reported in Table 4.

Table 5 was constructed listing each of the 380 entries. Each feature found in an item in common with the model was noted. Results are as follows:

The number of characteristics in common was plotted against the number of specimens sharing those characteristics. The plot gave roughly a bell-shaped curve.

Writings were next sorted by characteristics. A batch of writings was first sorted by looking at all writings starting with the most common characteristic. In this sample of hand printing it was the predominant right-hand slant, appearing 85.5% of the samples. Samples sorted by these criteria were then divided into smaller groups by using the next most frequent criteria, the lower loop of the capital "S" larger than the upper loop. This led to the following subdivisions reported in Table 6:

A numerical sorting scheme produced a better match with the model than was accomplished in

Table 4. Percentage of samples with observed features.

Feature	Number of Times Observed	Percentage of Samples with Feature
A	133	35.0
B	125	32.9
C	322	84.7
D	243	63.9
E	210	55.3
F	325	85.5
G	112	29.5
H	95	25.0
Ave.	195.6	51.5

Table 5. Observations of features in common.

Features in Common	**Number of Times Observed**
8	1
7	11
6	38
5	78
4	99
3	77
2	54
1	19
0	1

Table 6. Number of common characteristics and frequency of combination occurrence.

Sorting Criteria	Probability of Combination Occurrence	Number of Times Combination Observed	Frequency of Combination Occurrence
F	$(½)^1 = 0.50$	325	1/1.6 (0.625)
FC	$(½)^2 = 0.25$	189	1/2 (0.500)
FCD	$(½)^3 = 0.125$	137	1/3 (0.333)
FCDE	$(½)^4 = 0.0625$	78	1/5 (0.200)
FCDEA	$(½)^5 = 0.0312$	40	1/9.5 (0.105)
FCDEAB	$(½)^6 = 0.0156$	13	1/29 (0.034)
FCDEABG	$(½)^7 = 0.0078$	6	1/63 (0.016)
FCDEABGH	$(½)^8 = 0.0039$	1	1/380 (0.003)

the cursive writing exercise. The printed model had more uniform features from one duplicate to the next.

Statistical Analysis

The next part of this study dealt with a study of the correlation, if any, between different features. Recent studies have looked at feature distributions but not correlation of one feature to another (Johnson, 2017). Chi square (X^2) is the simplest statistical analysis of the frequency and independence of joint characteristics. This kind of information is useful to the document examiner who does not want to spend time evaluating a feature if it does not provide additional information, i.e. is not an independent characteristic. On the other hand, if two features that in the population are highly dependent diverge in both questioned and known documents, this evidence can help provide evidence of common origin (Ansell, 1979).

Chi square is used to determine the statistical reliability of departures in the data from the theoretical prediction based on the assumption of independence. From the frequency value of each characteristic as determined by an actual sample count, the correlation between any two features was first computed by multiplying the two frequencies to give an expected value and to then compare this value with the observed value which was obtained by sorting through the samples one more time, this time doing an actual count of those writings that had both, one but not the other, and neither characteristic. An example of a chi square calculation is given below. The example is to determine if a correlation exists between (E) descender of "f" and "p" of nearly equal length and (G) height of "l" less than or equal to the height of "d" and if there is a correlation between these two characteristics in the cursive writing sample. The actual count is noted with the expected count in parentheses.

The null hypothesis, H_0 states that the proportion of writings falling into each of the categories with the noted characteristics. If the null hypothesis is true, then there should be little chance of deviation from the expected frequency which assumes a correlation between the two classes of characteristics noted. If there is a large discrepancy between observed and expected frequencies, then the chi square value will be large indicating that it is less likely that the observed frequencies came from the population on which the null hypothesis and expected frequencies are based. In other words, if there is a large X^2 then the two characteristics are not correlated. The chi square formula is:

	Descender of "f" and "p" of equal length	Other
Height of "l" less than or equal to height of "d"	373 (377)	537 (556)
Other	67 (8)	131 (118)

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

O = the frequencies observed

E = the frequencies expected

$\sum$ = the 'sum of'

The degrees of freedom (d.f.) are the total number of categories minus one which in this case is d.f. = 3. Chi square value at 95% confidence level is obtained from tabulated values in any standard statistics text (Siegel, 1988).

For d.f. = 3, a X^2 at the 95% significance level, X^2 (.05) = 7.82. Since the X^2 is less than the variation that one might expect to see by chance, there exist a correlation between factors E and G. In other words, long descenders in certain letters tend to go with long ascenders in certain letters within the same writing. Since these traits tend to be correlated with each other, they would not be as suitable as traits wholly independent in identifying common writers or distinguishing between two or more writers. No correlations were found between the following features that were subjected to chi square calculations in the cursive writing sample:

1. (A) pen lift between "g" and "f" and (C) connected, cursive "S."
2. (D) loop of "g" longer than loop of "d" or "f" and (E) descender of "f" and "p" of nearly equal length.
3. (B) initial spike of "r" is higher on the left side and (L) apex of "r" equal or higher than top of the "p."
4. (H) overall right-hand slant and (I) open loop in descender of "p."

No correlations were found between the following calculated features in the hand printed samples:

1. (B) capital "A" with single sharp angle at apex and (H) initial stroke of "A" separate from rest of body.
2. (A) majority of letters are capitals and (E) equal spacing of letters.

Conclusions and Further Study

The ability to cut writing samples by roughly half with the addition of each new characteristics emphasizes the point that a combination of features may be rare even though each of the features making up the combination is itself common. Characteristics that I thought were unique letter forms were found to be used by other writers. An example is the unusual capital "S" found in cursive samples 61, 374, 464, 831, 868 and 918 (Figure 6). Hand printed samples provide just as much individuality as cursive samples and can be classified just as readily.

Because of the variability within the cursive writing model, not all 12 features were found in all 11 models. One of these models, picked at random and placed in the pool with the other1130 samples would not necessarily have been retrieved by a strictly numerical count of characteristics. Nicholson in his 1995 study of computer retrieval systems for handwriting recommends that at least 90% of characteristics between questioned and retrieved samples should be comparable.

The results of this project suggest other possible studies. It would be desirable to see how well other people can count the same characteristics and to see if they can retrieve the same model writing. However, taking counts of 12 character-

Table 7. Chi square analysis of features.

Observed (O)	Expected (E)	$(O-E)^2$	$(O-E)^2/E$
373	377	16	0.04
67	80	169	2.11
537	556	36	0.65
131	118	169	1.43
		Total	X^2 = 4.23

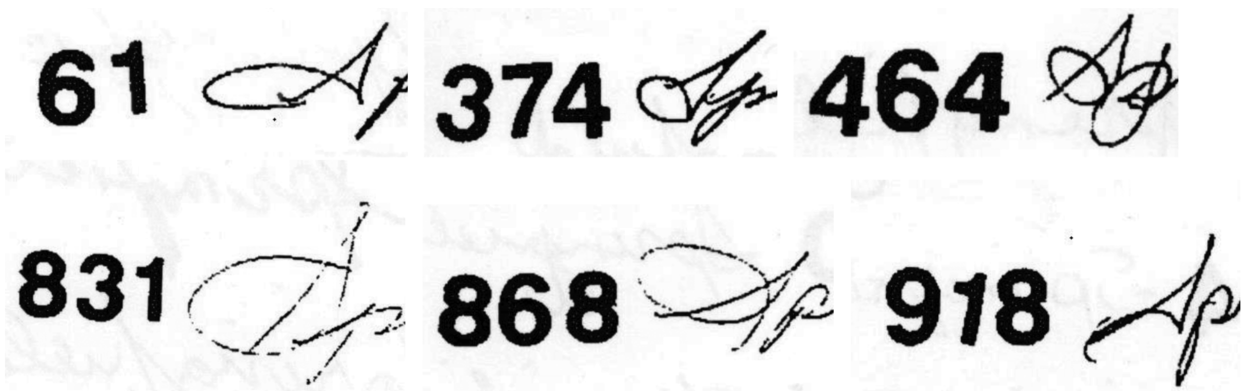


Figure 6. Illustrations of common letter formations between writers.

istics in a pool of over 1100 writers requires about 15 hours. The entire set of writings is available and might be included as part of a training exercise for beginning document examiners to reinforce the value of combined features in discriminating between writers with similar handwriting.

Another possible project might involve a numerical ranking of the writing samples. It might be possible to rank writings from closest in appearance to the model to the least similar and in the process formulate a ranking of opinions from "most probable" to "can be eliminated." Can this data be used to evaluate the concept of partial identity? At what level do the similarities approach those observed by chance in the general population? What are the minimum number of features needed for a meaningful comparison? Should such characteristics receive equal weight? What is the range of similarities suggesting an identification, elimination or inconclusive (Cashman, 1981)? Can a numerical system be developed for features such as line quality? The range of degrees of similarity to a model might be of some assistance in defining closeness of match using a scale of similar features (McAlexander, et al., 1991).

Moore, in his handwriting classification system (1945) suggested uniform vs. non- uniform when comparing differences in pressure between up stroke and down stroke. Further subdivision and the attachment of numerical values to such comparisons might be difficult.

Correlation of more of the 12 characteristics would be another project. A two-by-two comparison of 12 features would involve (12 x 11)/2 or 66 comparisons.

References

Ansell, M. (1979). Handwriting classification in forensic science. *Visible Language, 13,* 239-251.

Baxter, P.G. (1973). Classification and measurement in forensic handwriting comparisons. *Medicine, Science, and the Law, 13* (3), 166-184.

Britt, S.H. & Mensh, I.N. (1943). The identification of one's own handwriting. *Journal of Criminal Law and Criminology, 34,* 50-60.

Cashman, P.J. (1981). *The utilization of similarity analysis in comparison identification.* Unpublished doctoral dissertation, University of California, Berkeley.

Chabot, C. (1871). *The handwriting of Junius.* London: William Clawson & Sons.

Eldridge, M.A., Nimmo-Smith, I., Wing, A.M., & Totty, R.N. (1984). The variability of selected features in cursive handwriting: categorical measures. *Forensic Science Society Journal, 24* (3), 179-219.

Freeman, F.N. (1918). The handwriting movement: A study of the motor factors of excellence in penmanship. *Supplementary Education Monographs vol. 2,* no. 3 (August), 1-169.

Hall, M.G. & Hardcastle, R.A. (n.d.). Is a universal handwriting specimen feasible? *HOCRE Report No. 8-19/43.*

Hardcastle, R.A., Thornton, D., Totty, R.N. (1986). A computer-based system for the classification of handwriting on cheques. *Forensic Science Society Journal, 26* (6), 383-392.

Editor's note: Additional "Springfield" images can be downloaded at http://www.asqde.org.journal/journal.html

Harris, J.J. (1958). How much do people write alike? *Journal of Criminal Law, Criminology and Police Science, 48*, 647-651.

Harrison, W.R. (1966). *Suspect documents.* London: Sweet & Maxwell.

Havelock, E.A. (1976). *Origins of Western literacy (monograph series no. 14.* Toronto: The Ontario Institute for Studies in Education.

Hilton, O. (1983). How individual are writing habits? *Journal of Forensic Science, 28* (3), 683-685.

Hilton, O. (1982). *Scientific examination of questioned documents.* New York: Elsevier.

Hilton, O. (1958). The relationship of mathematical probability to the handwriting identification problem. *Royal Canadian Mounted Police Seminar, no. 5,* 121-128.

Johnson, M.E., Vastrick, T.W., Boulanger, M., & Schuetzner, E. (2017). Measuring the frequency of handwriting and handprinting characteristics. *Journal of Forensic Sciences, 62* (1), 142-163.

Kam, M., Wefstein, J., & Conn, B. (1994). Proficiency of professional document examiners in writer identification. *Journal of Forensic Sciences, 39* (1), 5-14.

Lawton, S.B., Currie, A.B. (1980). *Handwriting I. Instruction in handwriting in Ontario schools. II. Handwriting: An annotated bibliography.* Toronto, Ontario: Ontario Government Bookstore.

Livingston, O.B. (1959). A handwriting and pen-printing classification system for identifying law violators. *Journal of Criminal Law, Criminology and Police Science, 49,* 487-506.

Lee, C.D. & Abbey, R.A. (1922). *Classification and identification of handwriting.* New York: D. Appleton & Co.

McAlexander, T.V., Beck, J. & Dick, R. (1991). The standardization of handwriting opinion terminology. *Journal of Forensic Science, 36* (2), 311-319.

Moore, J. (1945). Handwriting classification: A description of the system employed by the Nottingham City Police. *The Police Journal (England), 18* (1), 39-61.

Muehlberger, R.J., Newman, K.W., Regent, J., & Wichman, J.G. (1977). A statistical examination of selected handwriting characteristics. *Journal of Forensic Science, 22* (1), 206-215.

Newland, T.E. (1932). An analytical study of the development of illegibilities in handwriting from the lower grades to adulthood. *Journal of Education Research, 26* (4), 249-258.

Nicholson, P.J. (1995). A system for the classification of block capital handwriting. Paper presented at the meeting of the American Society of Questioned Document Examiners, Chicago, IL.

Purtell, D.J. (1980). Modern handwriting instructions, systems and techniques. *Journal of Police Science and Administration, 8* (1), 66-68.

Rhodes, E.F. (1978). *The implication of kinesthetic factors in forensic handwriting comparisons.* Unpublished doctoral dissertation, University of California, Berkeley.

Siegel, S. & Castellan, N.J. (1988). *Non-parametric statistics for the behavioral sciences.* New York: McGraw-Hill, Inc.

Srihari, S.N. (2013). *Statistical examination of handwriting using automated tools.* Washington, D.C.: U.S. Department of Justice, National Institute of Justice. Document 241743.

Tan-Lin, A. (1981). *An investigation into the development course of preschool kindergarten. aged children's handwriting behavior.* Unpublished doctoral dissertation. Southern Illinois University, Carbondale.

Templin, E.M. (1958). *A comparative study of the legibility of handwriting of 454 adults trained in three handwriting styles: All cursive, all manuscript or manuscript-cursive.* Unpublished doctoral dissertation, New York University.

Wing, A.M. & Nimmo-Smith, I. (1987). The variability of cursive handwriting measure defined along a continuum: letter specificity. *Forensic Science Society Journal, 27* (5), 297-306.

Wing, A.M. (1979). Variability in handwritten characters. *Visible Language, 13,* 283-298.