J Forensic Sci, March 2007, Vol. 52, No. 2
doi:10.1111/j.1556-4029.2006.00355.x
Available online at: www.blackwell-synergy.com

TECHNICAL NOTE

Chi-Keung Li,[1] *Ph.D.; Chi-Ting Yang,*[2] *M.Phil.; Nai-Leung Poon,*[1] *Ph.D.; and Wing-Kam Fung,*[2] *Ph.D.*

Significance of Sequence of Strokes in Chinese Handwriting Examination*

ABSTRACT: This paper reports a statistical study on the sequence of strokes of 61 commonly encountered Chinese radicals and characters written by 372 invited subjects. The distribution of different writing sequence of these Chinese radicals and characters was examined. Comparison of the sequence of strokes executed by the subjects with the standard rule of writing these Chinese radicals and characters revealed that around 60% of the subjects wrote in the correct sequence. Pair comparison of sequences of strokes in Chinese handwriting among the 372 subjects was also performed. The results demonstrated that no two individuals wrote all the 61 radicals and characters with the same sequence of strokes. The findings indicate that, despite some basic rules governing the writing sequence of Chinese characters, writers tend to develop their own habits. The findings also support the hypothesis that the handwriting of experienced writers is individual.

KEYWORDS: forensic sciences, sequence of strokes, statistical study, Chinese handwriting, individuality in handwriting

Chinese characters, compared with English letters, are relatively complicated in their structures (1). Unlike English words, which are made up of linearly arranged discrete units (alphabet letters and characters), all Chinese characters are written within a framework of an "imaginary box" (2) with component radicals arranged in a two-dimensional array so that they can be presented neatly with similar dimensions for easy reading. A Chinese character is made up of one or more basic components called radicals. Each radical is formed by eight basic strokes namely vertical, horizontal, angular, slash, saber, hooked, tick, and dot as illustrated in Fig. 1. The sequence of strokes of a Chinese radical/character is governed by standard writing rules that state that all basic strokes, with the exception of the tick, are written in a descending manner from top to bottom, left to right, and periphery to center. Following these rules, not only are the strokes of a particular character written smoothly and rapidly (3) but they are presented in a manner that reflects harmony, beauty, and legibility (4). Despite the presence of standard rules governing the act of writing, very few writers and prominent calligraphers strictly follow these conventions.

Owing to the two-dimensional structure of the Chinese character, a structural relationship exists among certain neighboring strokes in the character that makes them easier to connect without affecting the formation or appearance of the character. As most of the Chinese characters are composed of multiple strokes, there are many possible chronological sequences for writing a Chinese character. As an example, four common ways of writing the simple Chinese character "王" (King), which consist of four strokes, are illustrated in Fig. 2.

Various methods have been reported for determining the stroke sequences, from the simple microscopic observation to the complicated instrumental method such as a scanning electronic microscope. In Chinese handwriting examination, as most of the normal courses of business writings are written in cursive style, the chronological stroke sequence can readily be determined by examining the start and end of strokes that are either joined by fine connecting extensions, or by observing the direction of the tapering end of one stroke and the flying start of another (5). Despite the fact that this does not apply to the regular style of handwriting in which the strokes are written separately, the stroke sequences can also be determined by specialized methods such as the Kromkote paper (6), burr striations (7), or the scanning electronic microscope (8).

As natural handwriting is the result of an individual's physical and mental endowment as influenced by condition at the time of writing, system, and the methods of instruction under which it is acquired, a writer tends to develop his/her peculiar writing sequences according to his/her needs and writing ability. Therefore,

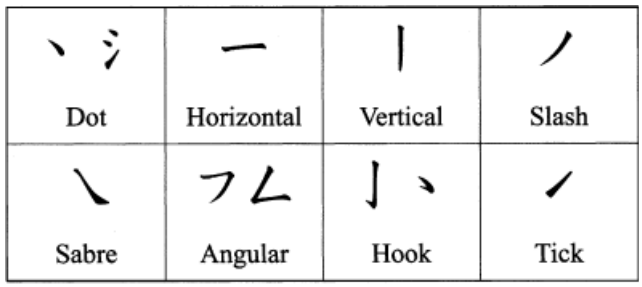

FIG. 1—*Eight basic strokes in Chinese handwriting.*

[1]Questioned Documents Section, Government Laboratory, Ho Man Tin Government Laboratory, 88 Chung Hau Street, Kowloon, Hong Kong Special Administrative Region, China.
[2]Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong Special Administrative Region, China.

*The results of this work have been presented in the International Association of Forensic Science 2005 Triennial Conference.

Received 6 Nov. 2005; and in revised form 9 July 2006; accepted 5 Aug. 2006; published 30 Jan. 2007.

Radical/ Character	Sequence of strokes	Choice	Percentage
王 King	一 二 三 王	A	11.0
	一 丅 干 王	B	22.3
	一 二 干 王	C	66.1
	丨 丅 干 王	D	0.5
	Other	E	0.0

FIG. 2—*Difference sequence of strokes of writing the Chinese character 王 (King) and their distribution.*

the structural relationship of strokes and the sequence in which they are executed are of particular significance for identifying or discriminating authorship in Chinese handwriting examination.

This paper reports a statistical study on the writing sequence of strokes of 61 commonly encountered Chinese radicals and characters written by 372 subjects. In this study, the different stroke sequences of these Chinese radicals and characters and the degree to which they conform with the standard rules of writing were examined. In addition, a preliminary study of the sequence of strokes used to write a radical when it is located in different Chinese characters was also performed. In order to quantify/demonstrate the individuality in handwriting sequence, pair comparison of 372 subjects was performed. A total of $372 \times 371/2 = 69{,}006$ pairs of comparison were computed for each Chinese radical and character, and the number of pairs with exactly the same writing sequence in the selected radicals and characters was counted. It is hoped that the results of this analysis can be used to demonstrate the homogeneity of the handwriting patterns as well to validate the hypothesis that the handwriting of experienced writer is individual. The findings in the analysis can provide a scientific argument to substantiate the hypothesis of individuality of handwriting as published scientific articles in this area are rare (9,10).

Materials and Methods

Sixty-one commonly encountered Chinese radicals/characters were selected in this study as listed in Table 1, and a questionnaire containing these selected radicals/characters with various stroke sequences choices, from two for the simple character "刀" up to 14 for the more complicated character "門" was designed for data collection. Figure 2 illustrates an example of the choices of stroke sequences when writing the radical/character "王" (King) in the questionnaire. A total of 372 subjects from the local population were invited to participate in the study. A complete set of questionnaire containing all the stroke sequence choices for the 61 Chinese radicals/characters was issued to each of the participants. They were asked to complete the questionnaire by selecting their choices of the writing sequence as exhibited in their normal course of business writing. In order to exhaust all the possibilities of writing sequences, those who wrote their radicals/characters other than the choices provided were requested to write down their own sequences of strokes in a blank area provided on the questionnaire. The background details of the subjects are documented in Table 2.

TABLE 1—*The 61 Chinese radicals and characters selected in this study.*

Number of strokes	Radicals/characters
2 to 3	阝刀力辶忄士也女山
4	礻開艹五王户戈手方水火爿片
5	乍主正用巨北甘世出生母田
6	兆糸 光共聿成米臣衣
7	里甬男我車
8	侖果妻長隹門青
> 8	禺垂重馬鳥樂

TABLE 2—*Personal details of the 372 invited subjects.*

Age		
10–14	56	15%
15–19	47	13%
20–24	19	5%
25–29	69	19%
30–34	61	16%
35–39	38	10%
40–44	39	10%
45–49	19	5%
50–54	15	4%
55 or above	9	2%
Sex		
Male	191	51%
Female	181	49%
Place of education		
HKSAR	339	91%
China	7	2%
Others	3	1%
HK/China	6	2%
HK/Others	17	4%
Education level		
Primary	14	4%
Secondary	132	35%
Postsecondary	55	15%
University	171	46%
Handwritten habit		
Left	5	1%
Right	367	99%
Frequency		
Often	207	56%
Sometimes	116	31%
Seldom	49	13%

Chinese character		Common structure
王 King	生 Live	王
女 Female	妻 Wife	女
里 Mile	重 Heavy	里

FIG. 3—*Three pairs of characters containing a similar structure.*

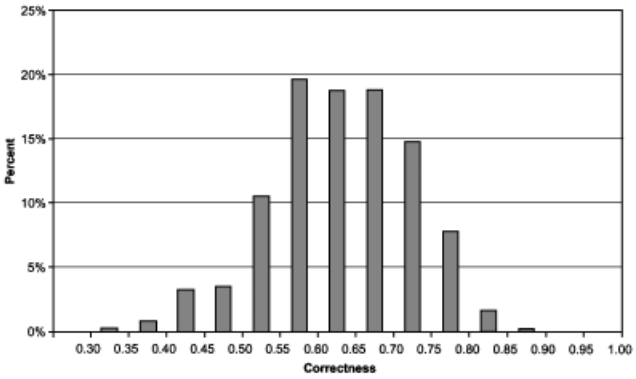

FIG. 4—*Distribution of the correctness of the sequence of writing of the 61 Chinese radicals and characters of the 372 subjects.*

Descriptive statistics for each variable were obtained. The deviation from the standard rule of writing of the 61 selected Chinese radicals and characters by the 372 subjects from the standard rule of Chinese characters writing was examined. In order to validate the individuality of handwriting, pair comparison of sequences of strokes in Chinese handwriting among the 372 subjects was also performed.

Among the 61 Chinese radicals and characters, 14 radicals/characters pairs embodied similar structures. Examples of these pairs are illustrated in Fig. 3. A preliminary study of the sequential association among these pairs was studied in our study in order to verify the dependency of sequence of strokes between characters sharing common radicals. A statistical test for independence or no

TABLE 3—*Sequence of strokes of writing the 61 Chinese radicals and characters used by the majority of the 372 invited subjects.*

Radicals/ Character	Sequence of Writing of the Majority	Percentage
1. 阝 **Ear**		51.1
2. 刀 Knife		97.0 *
3. 力 Force		95.4 *
4. 辶 Boat		99.2 *
5. 忄 Heart		58.9
6. 士 Scholar		95.2 *
7. 也 **Also**		90.6 *
8. 女 Female		66.7 *
9. 山 Hill		50.5 *
10. 礻 Divine		97.0 *
11. 開 **Open**		97.3 *
12. ⺾ Grass		43.5
13. 五 Five		50.3
14. 王 King		66.1 *
15. 户 **Family**		97.8 *
16. 戈 **Spear**		85.2 *
17. 手 Hand		99.5 *
18. 方 Square		98.4 *
19. 水 Water		94.4 *
20. 火 Fire		61.0 *
21. 爿 **Wood plank**		59.7 *
22. 片 Piece		62.1 *
23. 乍 **Suddenly**		81.7 *
24. 主 Host		63.7 *
25. 正 Positive		94.4 *
26. 用 **Use**		99.2 *
27. 巨 Big		48.1 *
28. 北 North		43.3
29. 甘 **Sweet**		52.2 *
30. 世 **Era**		50.3 *
31. 出 Exit		41.9
32. 生 Live		49.2
33. 母 Mother		65.6 *
34. 田 Field		86.8 *
35. 兆 Trillion		84.1 *
36. 糸 **Silk**		84.4 *
37. 光 Light		49.7 *
38. 共 Common		82.5 *
39. 聿 **Pen**		77.7 *
40. 成 **Succeed**		83.3 *
41. 米 Rice		86.3 *
42. 臣 **Official**		34.1
43. 衣 Coat		88.2 *
44. 里 Mile		60.2
45. 甬 **Drum stick**		97.5 *
46. 男 Male		29.8
47. 我 Me		86.8 *
48. 侖 **Arrange**		95.2 *
49. 車 Car		86.0 *
50. 果 Fruit		86.6 *
51. 妻 Wife		76.1 *
52. 長 Long		68.8 *
53. 隹 Bird		38.7
54. 門 Door		29.8
55. 青 Green		44.9
56. 禺 **District**		63.4 *
57. 垂 **bend down**		27.2 *
58. 重 Heavy		32.5
59. 樂 Joy		80.6
60. 馬 Horse		51.6
61. 鳥 Bird		63.7 *

*The sequence of strokes agrees with the rules of writing.

association can be constructed as follows: consider that we are interested in testing among subjects whether there is an association in the sequences of strokes of writing two Chinese characters A and B. Suppose there are R and C ways of stroke sequences for writing the characters A and B, respectively (e.g., there are four ways of writing the character "王" [king] as shown in Fig. 2). Let f_{ij} be the observed number of subjects (among 372 subjects) who have used the ith way in writing character A and jth way in writing character B. In testing the hypotheses

H_0: written sequences for the two Chinese characters are independent, versus

H_1: written sequences for the 2 Chinese characters are dependent,

the following test statistic has been constructed (11):

$$\chi^2 = \sum_{i=1}^{R}\sum_{j=1}^{C}\frac{(f_{ij}-e_{ij})^2}{e_{ij}}$$

where e_{ij} is the corresponding expected number of subjects in using the ith and jth ways in writing characters A and B, respectively, when the null hypothesis H_0 holds. The test is distributed as a χ^2 distribution with $(R-1)(C-1)$ degrees of freedom under H_0, and the probability value (p-value) of the test statistic can be obtained accordingly. If the p-value is smaller than the 5% level of significance, we reject the null hypothesis H_0 and may conclude that among subjects the written sequences for the two Chinese characters are dependent/associated.

The above-mentioned statistical analyses including preliminary descriptive statistics, independence testing as well processing of data for pair comparison were conducted with the aid of SPSS for Windows platform (12).

Results and Discussion

The personal details of the 372 invited subjects are listed in Table 2. They were subdivided into 10 age groups ranging from 10–14 to 55 or above. The age groups 25–44, being the key contributors of the samples, account for 55% of the total. The majority of the subjects received their education in Hong Kong (91%) and nearly half of them (46%) had university–level education. About 87% of the total subjects indicated that they were frequent writers.

In the present study, in order to avoid the misinterpretation of the sequence of strokes of writing, the 372 subjects were asked to select their own writing sequence from the questionnaire provided. The stroke sequences used by the subjects to write the 61 Chinese radicals and characters were compared with the standard writing sequences. The overall average value of "correctness" is about 63% (SD = 9%), and the correctness for individual subjects ranged from 33% to 85% as illustrated in Fig. 4. About 8% of the subjects wrote <50% of the radicals/characters in the correct sequences whereas about 10% of the subjects wrote more than 45 radicals/characters in the correct manner. The stroke sequences of the 61 Chinese radicals and characters executed by the majority of the subjects are listed in Table 3. The majority of participants wrote 45 radicals in the correct sequence and, of this number, 26 radicals were written correctly by over 80% of the writers who participated in the study.

Among the 61 selected Chinese radicals and characters, 14 pairs possess a similar structure (Fig. 5 illustrates the similar stroke sequences for the writing for the characters "里" [Mile] and "重" [Heavy] in which both contained the structure "里"). In order to demonstrate whether the subjects wrote a similar structure with similar writing sequences, a statistical χ^2-test for independency was performed to investigate the association of writing sequence among the radicals/characters pairs. The results of the statistical tests are summarized in Table 4. As the test statistics of all pairs yielded a p-value less than the 0.05 level of significance, the significance of sequential association among the paired characters is strongly suggested. The preliminary study indicates that the majority of the subjects tended to adopt the same stroke sequence for the same structure (radical) located in different characters. A detailed study will be conducted on Chinese characters with more complicated structures.

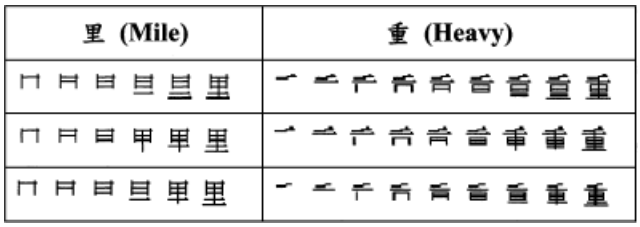

FIG. 5—*Similar sequences of writing for the characters 里 (Mile) and 重 (Heavy).*

A pair-wise comparison of the 372 subjects was performed to demonstrate the individuality of Chinese handwriting. Table 5 gives the findings of two pair-wise comparisons, from the simplest to the most complicated and vice versa in the reverse direction. Referring to the set of values in normal font at the top of Table 5, a comparison of the first five simple Chinese radicals with two to three strokes shows that about 21% of all 69,006 individual pairs share exactly the same written sequence. When 10 radicals/characters were compared, the percentage of pairs with exactly the same sequence of writing for the 10 radicals/characters declined dramatically to 3.4%. There is no one pair of subjects sharing exactly the same writing sequence when (the first) 30 or more radicals/characters were considered simultaneously.

With respect to the other set of values in Table 5 in *italics* font at the bottom of Table 5, a comparison of the last five complicated

TABLE 4—*p-Values of χ^2-test for independence for the 14 pairs of Chinese radicals/characters.*

First Character	Second Character	P value
士	王	0.014
士	主	0.007
士	生	0.016
王	主	0.000
王	生	0.000
主	生	0.000
女	妻	0.000
山	出	0.000
巨	臣	0.000
北	兆	0.010
聿	里	0.000
聿	車	0.000
聿	重	0.024
里	重	0.000

TABLE 5—*Pair-wise comparison of the sequence of stokes of writing of the Chinese radicals and characters of the 372 invited subjects; from the simplest character to the most complicated and vice versa in a reverse order (starting from the last character 61).*

Characters in difference	Characters for comparison							
	First 5	First 10	First 20	First 30	First 40	First 50	All 61	
0	21.465	3.395	0.148	0	0	0	0	
1	46.173	16.194	0.935	0.004	0	0	0	
2	27.516	30.250	3.979	0.074	0.004	0	0	
3	3.950	28.750	10.486	0.317	0.007	0.001	0	
4	0.884	15.200	17.375	0.961	0.070	0.007	0	
5	0.012	5.008	21.043	2.400	0.199	0.030	0	
6		1.068	19.391	4.762	0.543	0.070	0.001	
7		0.128	13.744	8.317	1.129	0.184	0.003	
8		0.007	7.779	11.782	2.277	0.432	0.006	
9			3.469	14.205	3.898	0.852	0.022	
10			1.268	15.332	5.776	1.578	0.046	
. . .			0.314	14.255	8.015	2.695	0.122	
. . .			0.062	11.276	9.672	3.885	0.194	
. . .			0.006	7.956	11.008	5.563	0.383	
				4.549	11.627	7.228	0.662	
				2.371	11.143	8.549	1.122	
				1.020	9.976	9.437	1.742	
				0.313	8.173	9.853	2.493	
	0.003			0.090	6.160	9.646	3.459	
	0.010			0.014	4.368	8.641	4.607	
	0.025			0.003	2.755	7.675	5.715	
	0.065				1.665	6.259	6.996	
	0.122	*0.001*			0.852	5.008	7.825	
	0.239	*0.003*			0.410	3.908	8.463	
	0.372	*0.012*			0.187	2.872	8.543	
	0.612	*0.052*			0.068	2.153	8.085	
	0.927	*0.125*			0.017	1.393	7.665	
	1.391	*0.271*			0.001	0.946	6.928	
	1.745	*0.509*	*0.006*			0.526	5.692	
	2.297	*0.801*	*0.028*			0.322	4.692	
	2.945	*1.249*	*0.087*			0.178	3.781	
	3.781	*1.658*	*0.256*			0.074	2.945	
	4.692	*2.324*	*0.494*			0.022	2.297	
	5.692	*3.077*	*0.783*			0.007	1.745	
	6.928	*4.043*	*1.197*			0.004	1.391	
	7.665	*5.244*	*1.778*	*0.003*			0.927	
	8.085	*6.275*	*2.406*	*0.058*			0.612	
	8.543	*7.895*	*3.404*	*0.193*			0.372	
	8.463	*8.727*	*4.639*	*0.520*			0.239	
	7.825	*9.159*	*6.360*	*0.940*			0.122	
	6.996	*9.119*	*7.678*	*1.525*			0.065	
	5.715	*8.906*	*9.295*	*2.116*			0.025	
	4.607	*7.994*	*10.193*	*2.937*	*0.001*		0.010	
	3.459	*6.759*	*10.263*	*4.397*	*0.042*		0.003	
	2.493	*5.095*	*9.961*	*6.650*	*0.317*			
	1.742	*3.946*	*8.856*	*8.782*	*0.965*			
	1.122	*2.730*	*7.220*	*10.806*	*2.106*			
	0.662	*1.787*	*5.571*	*12.340*	*3.901*			
	0.383	*1.045*	*3.950*	*12.631*	*6.688*			. . .
	0.194	*0.652*	*2.545*	*11.492*	*10.627*			. . .
	0.122	*0.293*	*1.477*	*9.227*	*14.952*			. . .
	0.046	*0.143*	*0.862*	*6.708*	*17.494*	*0.901*		*10*
	0.022	*0.070*	*0.393*	*4.327*	*16.613*	*6.160*		*9*
	0.006	*0.029*	*0.183*	*2.380*	*12.589*	*16.157*		*8*
	0.003	*0.006*	*0.080*	*1.196*	*7.767*	*24.995*		*7*
	0.001	*0.000*	*0.035*	*0.497*	*3.837*	*24.234*		*6*
	0	*0.001*	*0.000*	*0.190*	*1.507*	*16.448*	*7.483*	*5*
	0	*0*	*0.001*	*0.070*	*0.456*	*7.772*	*30.037*	*4*
	0	*0*	*0.000*	*0.016*	*0.119*	*2.739*	*36.568*	*3*
	0	*0*	*0.001*	*0.001*	*0.017*	*0.527*	*20.490*	*2*
	0	*0*	*0*	*0*	*0*	*0.064*	*4.946*	*1*
	0	*0*	*0*	*0*	*0*	*0.003*	*0.477*	*0*
Characters for comparison	All 61	Last 50	Last 40	Last 30	Last 20	Last 10	Last 5	Characters in difference

Percentages among 69,006 paired comparison for differences in characters for the forward comparison are given in normal fonts in the upper triangular part of the table, while percentages for the reverse order comparison are in italic fonts in the lower triangular past of the table.

Chinese characters revealed a great discrepancy from the previous set where only 0.48% of pairs' comparison show no differences in writing sequence. When the last 10 characters were considered, only 0.003% (two out of 69,006) of all individual pairs had the same sequence of writing of the characters. The proportion of homogeneity among the pair comparison decreased dramatically when more Chinese characters were taken into account. No pair shared exactly the same writing sequence when (the last) 20 or more radical/characters have been compared. The results obtained from two sets of pair-wise comparison indicated that the discrimination power increases with the complexity of the Chinese radicals/characters. In addition, the findings support the hypothesis that no two individuals wrote the same.

Conclusion

Handwriting examination is based on the hypothesis that natural handwriting, being a product of a long-term adaptation to the needs and ability of a person, will be *peculiar* to a particular person. Owing to these complexities, the probability of any two persons having exactly the same writing characteristics is extremely rare. Despite the presence of a great number of general similarities, *a single consistent difference* that cannot be logically accounted for and reasonably explained between two writings is a strong indication of different authorship.

The methodology used in Chinese handwriting examination is comparable to English handwriting. The line quality, fluency of writing, pen pressure variation, writing movement, and connection of strokes are significant attributes that could provide a clue to authorship (6). However, the first three qualities are hard to be accurately assessed or measured in an objective manner. On the other hand, stroke sequence can be a good means of discriminating documents written by different authors. Despite the standard rules governing the sequence of writing in Chinese handwriting, the two-dimensional arrangement of various components of a character can lead to a large number of possible sequences. The results in this study not only provide the statistical distribution of different stroke sequences of writing the same radical/character that can be used in the real case examination, but also demonstrate that an individual writer will develop his/her own sequence of writing habits that deviate from the standard method of writing.

Although stroke sequence is only one of the many characteristics that can be used to identify Chinese handwriting, the presence of consistent differences in stroke sequence between two writing specimens constitutes significant evidence of different authorship. Pair comparison on the sequences of strokes of 61 Chinese radicals/characters among 372 invited subjects has been introduced in our study. An analysis of the collected data demonstrates the individuality of Chinese handwriting. A much higher degree of discrimination could be achieved by considering other writing attributes in addition to sequence of strokes.

Acknowledgments

The authors would like to express their gratitude to Dr. T. L. Ting, the Government Chemist of Hong Kong, for his interest and encouragement of this work, and to Mr. S. C. Leung, Assistant Government Chemist, for his valuable advice on this work and for proofreading the paper. The work was partly supported by an HKU CRCG research grant.

References

1. Leung SC, Tsui CK, Cheung WL, Chung MWL. A comparative approach to the examination of Chinese handwriting—the Chinese character. J Forensic Sci Soc 1985;25:255–67.
2. Björkstén J. Learn to write Chinese characters. New Haven: Yale University Press, 1994.
3. Tan HP. What's in a Chinese character. Beijing: New World Press, 2002.
4. Fazzioli E. Understanding Chinese characters: a beginner's guide to the Chinese language. London: William Collins Sons & Co. Ltd., 1987.
5. Leung SC. The scientific examination of Chinese handwriting. Forensic Sci Rev 1994;6:97–145.
6. Igoe TJ, Reynolds BL. A lifting process for determining the writing sequence of two intersecting ballpoint pen strokes. Forensic Sci Int 1982; 20:201–5.
7. Hung PS, Leung SC. Some observations on the morphology of a ball-point pen stroke. Int J Forensic Doc Exam 1995;1(1):18–31.
8. Waeschle PA. Examination of line crossing by scanning electron microscope. J Forensic Sci 1979;24:569–78.
9. Srihari SN, Cha SH, Arora H, Lee S. Individuality of handwriting. J Forensic Sci 2002;47:1–17.
10. Li CK, Poon NL, Fung WK, Yang CT. Individuality of handwritten Arabic numerals in local population. J Forensic Sci 2005;50(1):185–91.
11. Conover WJ. Practical nonparametric statistics. 2nd ed. New York: John Wiley, 1980.
12. SPSS version 11.0.2 for Windows Platform; http://www.spss.com

Additional information and reprint requests:
Chi-Keung Li, Ph.D.
Questioned Documents Section
Government Laboratory
Ho Man Tin Government Laboratory
88 Chung Hau Street, Kowloon
Hong Kong Special Administrative Region
China
E-mail: ckli@govtlab.gov.hk