Journal of the American Society of Questioned Document Examiners, Inc.
Volume 12, Number 2, December 2009

Relative Difficulty of Freehand Simulation of Four Proportional Elements in Arabic Signatures

Abdulaziz Al-Musa Alkahtani and Andrew W. G. Platt

Department of Forensic Sciences, Faculty of Sciences, Staffordshire University, College Road, Stoke-on-Trent, ST4 2DE, UK

This article compares the success with which 901 Arabic writers were able to simulate 4 proportional elements—size, slant, baseline, and spacing—in 2 Arabic signatures. Significant differences in ease of simulation were found among all 4 elements: slant was significantly more successfully simulated than the other 3; baseline was significantly more successfully simulated than spacing or size; and spacing was significantly more successfully simulated than size. The order of ease of simulation was found to be slant > baseline > spacing > size. This similarity between Arabic and Roman script data raises the possibility that slant and baseline, and possibly all 4 elements, might show similar relative rankings in other cursive scripts.

Introduction

This study investigated the relative success with which fluent Arabic writers were able to simulate 4 proportional handwriting elements—size, spacing, baseline, and slant—in 2 Arabic signatures. The main purpose of this comparison was to discover whether any of these elements is significantly easier to simulate. Studying the simulation of others' handwriting is as close as we can come to a controlled study of forgery. The use of statistics to assess the probability that differences in the rankings of the elements represent real differences in the larger population (Al-Musa Alkahtani and Platt 2009) places these findings on a more scientific basis than the qualitative studies that have dominated the field in the past; and recently a statistical analysis of some aspects of Arabic signatures has been reported (Al Haddad and others 2009). These studies, which reflect the principles accumulated over several centuries, have served the practice of forensic document examination well; and the few scientific tests of the techniques and assumptions they recommend have supported them (Srihari and others 2002). However, in 1993, the United States Supreme Court ruled that expert testimony admitted in trials (including fingerprinting and handwriting analysis) must be backed by scientific testing of the theories on which the techniques are based, error rates of the techniques, peer reviews of the tests, and acceptability in the relevant scientific community (*Daubert v Merrell Dow Pharmaceuticals* 1993). This challenge to technical forensic methods offers an opportunity to set these techniques on a more scientific foundation; and, for this reason, statistical studies in the field of forensic document examination are now of greater interest than in the past.

The phenomenon of freehand simulation is an important focus of forensic document examination and dominates studies of forgery in the field, primarily because most forgeries are of this type. Signatures were thus chosen as the target to be simulated. The literature on forgery deals almost entirely with Roman script. Data from other scripts should shed light on any similarities and differences between forgeries in different scripts, in addition to contributing to the knowledge of practitioners of forensic document examination in those scripts.

At present, there is no generally accepted view among forensic document examiners (FDEs) about whether any of these 4 elements is easier to forge than the others in cursive Roman script. An assessment of the accuracy with which 106 high school students simulated a number of elements in English signatures and found slant to be the most effectively simulated element, with considerable variation in the ease of simulation

Figure 1. The 2 target Arabic signatures with the 3 writing units circled.

between the same elements in different environments (Scheutzner 2006). Leung et al. (1993), in a study of the simulation of letter-like abstract symbols, found that errors made in simulating all elements of proportion tend to exaggerate extremes; thus, long strokes are made even longer, small sizes are made even smaller, and so on. However, they did not report which elements are most likely to be imitated incorrectly. In a study which shared some of the aims of this one, Lee et al. (2006) compared the success with which 62 Singaporeans simulated 12 specific features of a cursive signature. They found that slant, baseline, and size would present medium difficulty to simulators, while the instance of spacing would present high difficulty.

The relative sizes of letters have been suggested to be frequently ignored by forgers, and mistakes in their sizes are often mentioned in the literature as a sign of forgery (Osborn 1929; Harris 1958; Harrison 1981; Hilton 1993; Masson 1996; Huber and Headrick 1999). By contrast with the reported difficulty of simulating relative letter sizes, the overall size of handwriting is likely to be imitated accurately if the forger knows the typical size of the target writing (Horan 1985), as participants did in this study. Thus, it might be predicted that, in this study, relative sizes of letters would cause some problems for simulators but overall size of the signatures would be simulated accurately.

The baseline alignment of writing relative to an imaginary or ruled line is highly characteristic of individual writers but also unconsciously written and, therefore, less likely to be well reproduced by the forger (Osborn 1929; Huber and Headrick 1999). Some individuals also align individual letters in special ways with relation to the baseline, and immature handwriting of children or very poorly educated adults may wander on both sides of the baseline. For that reason, we might expect simulations of alignment with the baseline to reflect the simulators' habits rather than those of the target writers and, therefore, to be poorly simulated.

The slant of various strokes relative to the handwriting baseline has also been regarded as a useful marker of forgery. Osborn (1929) noted that the slant of different strokes is often highly

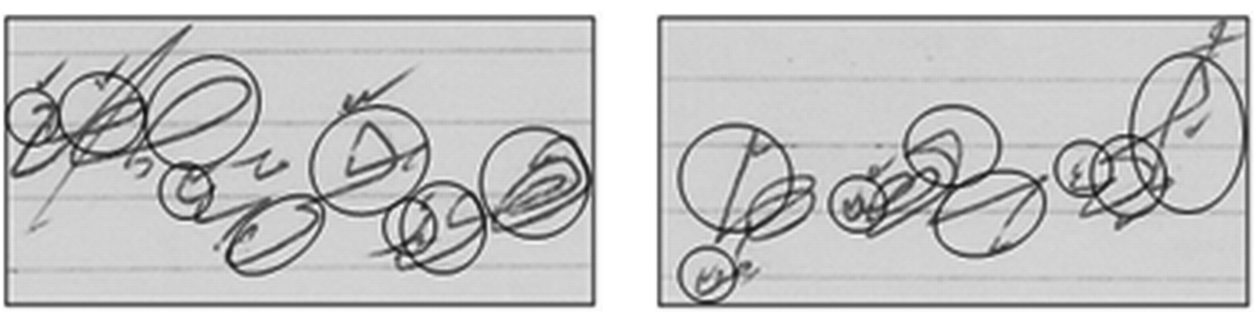


Figure 2. The 2 Arabic signatures with the size of letters marked with large, medium, and small circles respectively.

Figure 3. The 2 target signatures showing the baseline of the signature (dotted line) relative to the horizontal ruled baseline (solid line).

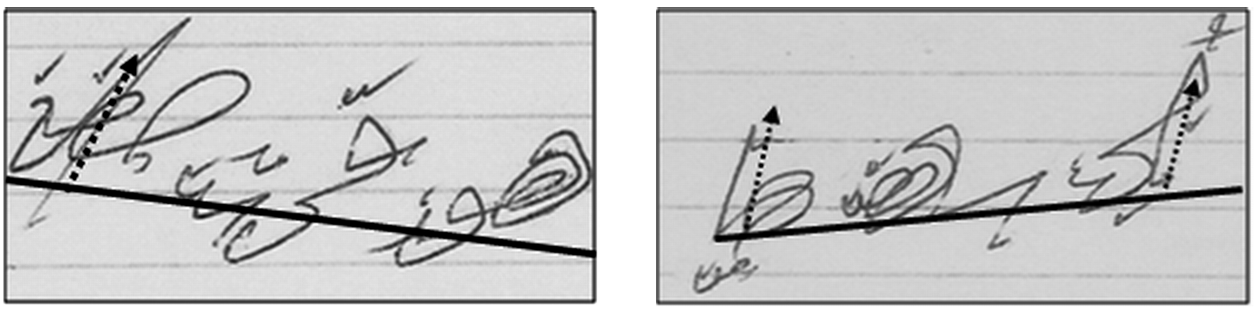
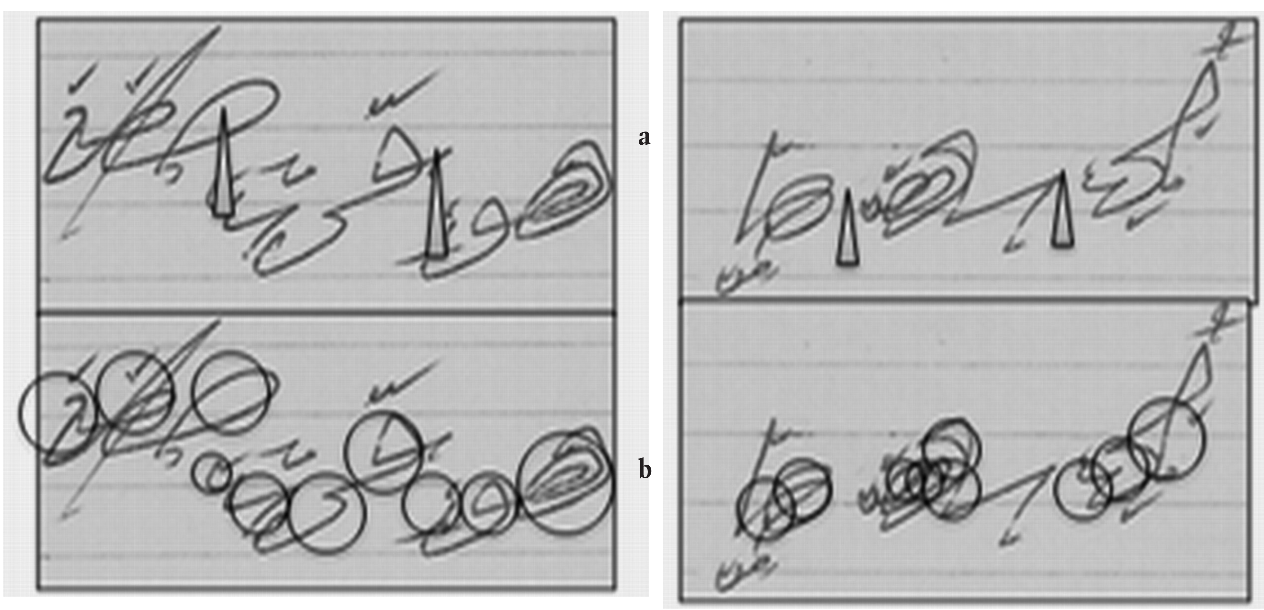


Figure 4. The target signatures showing the backward slanting strokes (dotted arrows); the rest of the letters are forward slanting.



Signature 1 Signature 2

Figures 5a and 5b. The triangles in the upper figures (a) mark the spaces between the 3 writing units of the signatures. For Signature 1 there are no spaces between the writing units, as the beginning and ending letters overlap. The circles in the lower figures (5b) mark the outer boundaries of the letters.

consistent in individual handwriting, and thus only a few degrees of difference between the slant in an original and in a forgery may be enough to identify fraud. Also, many people slant the up and down strokes of loops differently, which a simulator would not be likely to notice (Osborn 1929; Harris 1958; Harrison 1981; Hilton 1993; Masson 1996; Huber and Headrick 1999). From these observations, it would seem that slant is likely to be less well simulated than the other proportional elements.

The 4th proportional element, spacing, refers here both to spacing between writing units and to spacing between letters within writing units. Spacing is less frequently discussed as a sign of forgery in Roman script than the other 3 elements. Strongly slanted letters may allow little space between the letters and also between writing units. In fact, slanting letters may overlap onto each other's spaces, and slanted beginning and ending letters of writing units may also cause the writing units to overlap. By contrast, letters that are more upright are likely to be more compact, with little space between them, unless they have unusually long connections (Osborn 1929; Huber and Headrick 1999). It is not clear from the literature on forgery in Roman script how well spacing might be expected to be simulated.

Methods and Materials

The 2 Arabic signatures used as simulation models in this study are shown in Figure 1. The signatures include the potential for different errors in the same proportional element to be made. For example, both backward and forward slants are present, so that whatever the simulator's own ingrained habits of slant are, in at least 1 instance he or she will be forced to write with the opposite slant during the production of the simulated signature.

Size refers both to the relative sizes of different letters to each other and also to the overall size of the entire signature. Figure 2 shows the signatures with the unusually large, medium, and small letters marked by appropriately sized circles.

Baseline in this study refers to the alignment of letters or of total handwriting relative to a ruled or imaginary line. In many cases, the baseline of writing gradually rises higher above the ruled or imaginary baseline over the course of the signature or, less often, gradually falls farther below it; both tendencies are presented in the model signatures in this study, as shown in Figure 3.

Slant. Figure 4 shows the variety of slants in the model signatures and indicates the features which slant both with and against the overall slope of the writing baseline of the signature.

Spacing. In Figure 5a, the triangles on both signatures indicate the gap between writing units. For Signature 1, it can be seen that there are no spaces between the writing units, the last letters of the writing units overlapping with the 1st letters of the next; whilst in signature 2 there are clear gaps between the writing units. In Figure 5b, each letter in the 2 signatures is circled. The spaces between letters in Signature 1 are wide, with long connecting lines. Signature 2, by contrast, has rather narrow, cramped letters separated by very short connecting lines. Thus, Signature 1 has more widely spaced letters but more closely spaced writing units than Signature 2.

Data Collection

In order to collect attempts at simulation that were typical of the Arabic-writing population, simulations were elicited from a sample of 901 native Arabic writers. These people were recruited as members or employees of institutions, factories, schools, and businesses. The aim was to gather a representative cross-section of Arabic-speaking citizens in terms of gender balance, education, age, and occupation.

The participants were seated comfortably at tables or desks. Each was given a standard form attached to a clipboard, providing the same writing surface to participants, and asked to complete the form using a fine-point, 0.7 mm, black or blue ballpoint pen. The participants were allowed to study the target signatures for about half an hour.

The signatures consisted of 3 writing units each (Abjad Saa'fas Kaliman and Hawwaz Qurshat Dathagh), chosen to present as many letters as possible in as many word positions as possible. They are stylized so as to present different directions of slant and baseline and different patterns of spacing and size. This variety of expression of each of the 4 elements was included in the 2 signatures so that simulators could not duplicate any element correctly every time simply by writing according to their own ingrained habits.

The Ranking of Simulation Quality and Analysis of the Data

The rankings of the quality of simulations of the 4 elements for the 901 individuals were assessed

Element	Mean Rank
Slant	3.00
Baseline	2.70
Spacing	2.31
Size	1.99
N = 901	

Table 1. Friedman Test: Mean Ranks of the 4 Proportional Elements

Element	z	p	r
Slant Size	-18.604	< 0.0005	0.62
Slant Spacing	-14.382	< 0.0005	0.48
Size Baseline	-14.377	< 0.0005	0.48
Spacing Baseline	-9.159	< 0.0005	0.31
Spacing Size	-8.747	< 0.0005	0.29
Baseline Slant	-7.660	< 0.0005	0.26

Table 2. Wilcoxon Signed Rank Test Results

by a panel of 3 certified Arabic-writing FDEs. Each FDE assigned a numerical rank from 1 to 4 to each of the 4 elements—size, spacing, baseline, and slant—for each individual. The definitions of the 4 ranks are as follows.

Rank 1: Completely different
Rank 2: Significantly different
Rank 3: Similar
Rank 4: Identical

The rank of each element was based on every occurrence in the simulations of the 2 signatures, and the judging was carried out according to standard forensic document examination methods. The 2 judges' ranks were then averaged to give the final rank for each element for each individual. In the case of a fractional mean rank, the 3rd FDE rounded the number up or down, depending on his judgment of the simulation quality. Significance of the differences in the mean rankings of the 4 elements was assessed by a Friedman test with a significance level set at 0.05. Post hoc Wilcoxon signed ranks tests were carried out between each pair of rankings to determine which of the elements differed in the quality of simulation. A Bonferonni adjustment was made to the level of significance for these comparisons, this was set to 0.05/6 = 0.008. Following this, the size effect, r, was calculated to assess the degree to which the located differences exist (Cohen 1988).

Results and Discussion

The results of the Friedman test showed that there was a statistically significant difference in the quality of simulations across the size, spacing, baseline, and slant elements, χ^2 (3, n = 901) = 605.705, p <0.0005. Table 1 shows the mean ranks for size, spacing, baseline, and slant. The order of success in simulation was slant > baseline > spacing > size. Whilst the Friedman test showed that significant differences existed in the data, it did not tell us where these differences lay. To further explore this, we carried out a series of Wilcoxon Signed Rank tests between the rankings of pairs of the elements. For multiple comparisons of this nature. there is an increased possibility of Type 1 errors occurring; i.e., rejecting a null hypothesis (that there is no difference in the rankings) when it is actually true. To guard against this, the significance level at which the null hypothesis can be rejected is reduced by applying a Bonferroni adjustment. As there are 6 possible comparisons, the significance level for the Wilcoxon Signed Rank Tests was reduced to 0.05 / 6 = 0.008.

The results are shown in Table 2 and indicate that significant differences exist in all the pairs of rankings at a level <0.0005; i.e., well below the adjusted significance level of 0.008 for these tests.

For the large sample sizes used here, it is possible to detect statistically significant differences which are of relatively little importance. The effect size, r, provides a method of estimating this for every instance where significant differences have been found. This was estimated by calculating $r = |z/\sqrt{N}|$ with values of r > 0.5 representing a large effect, 0.5 > r >0.1 a medium effect size and r < 0.1 a small effect (Cohen 1988).

The effect sizes fall into 2 broad categories. Large effect sizes are observed for slant vs. size, slant vs. spacing, and baseline vs. size. Medium-sized effects are found for comparisons of baseline vs. spacing, baseline vs. slant, and spacing vs. size. It can thus be predicted that in forgeries of Arabic signatures, the slant will be simulated very much better than the size and spacing elements and, similarly, that the alignment of the signature to a baseline will be better reproduced than the overall size.

The relative ease with which these simulations were accomplished is not completely in accord with the conclusions from previous analysis of Roman script. The reasons for these differences were not immediately evident, however. Although the study by Scheutzner (2006) is not completely comparable to this study, her data agreed with these findings in showing that slant was considerably more successfully simulated than baseline. She did not present comparable information on spacing and size. Thus, her study of Roman script appeared to accord with some of the findings of this study of Arabic script, suggesting that the greater ease of simulation of slant than of baseline, at least, may be found not only in these 2 scripts but perhaps across other cursive scripts as well. Scheutzner also found that the same elements may vary widely in how easily they are simulated, depending on their environments. For example, size (length and width) of ascender loops is much better simulated than size of descender loops.

Conclusion

This study has shown that the 4 proportional elements were significantly different from each other in ease of simulation in Arabic script, and the similarities with Schneutzner's (2006) data on Roman script raise the possibility that slant is more easily simulated than baseline across many cursive scripts. It would be interesting to find comparable data for all 4 proportional elements in Roman script. In examining Arabic signatures, the relative sizes of letters and their spacing in comparison with the genuine signatures would seem to provide the best indicator of forgery.

Acknowledgments

We are grateful to the following FDEs for their time and effort in analyzing the simulated signatures: Ahmad Assiry, Jubran Qushaish, and Bandar Al-Hokbani.

Abdulaziz Al-Musa Alkahtani is grateful to the Counterfeiting and Forgery Department, Public Security, Ministry of Interior, Saudi Arabia, for financial support.

References

Al-Haddad, A.; White, P.C. and Cole M.D. (2009). Examination of a Collection of Arabic Signatures. *Journal of the American Society of Questioned Document Examiners.* Vol. 12, No. 1, pp. 35-53.

Al-Musa Alkhatani A. and Platt, A.W.G. A Statistical Study of the Relative Difficulty of Freehand Simulation of Form, Proportion and Line Quality in Arabic Signatures. *Science and Justice* (in press).

Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences,* 2nd ed. New Jersey: Lawrence Erblaum Associates.

Daubert v. Merrell Dow Pharmaceuticals (1993). 509 U.S. 579.

Epstein, G. (1987). Examination of the Josef Mengele Handwriting. *Journal of Forensic Sciences, Vol.* 32, No. 1, p. 100.

Harris, John J. (1958). How Much People Write Alike: A Study of Signatures. *Journal of Criminal Law, Criminology and Police Science*, Vol. 48, No. 6, pp. 647.

Harrison, W.R. (1981). *Suspect Documents: Their Scientific Examination*. Chicago: Nelson-Hall, Inc.

Hilton, O. (1993). *Scientific Examination of Questioned Documents*, revised ed. New York: Elsevier North Holland, Inc.

Horan, J.J. (1985). The Size of Forged Signatures. Presented at the meeting of the American Society of Questioned Document Examiners, Chicago.

Huber, R.A. and Headrick, A.M. (1999). Handwriting Identification: Facts and Fundamentals. Schaumburg, IL: CRCC Press LLC.

Lee, G-K.; Yup, B-S.; Yang, C-Y.; Lee, L-T.; Tan, S-K. and Tan, K-P. (2006). A Study on the Levels of Difficulty in the Simulation of Individual Characteristics in a Signature. *Journal of the American Society of Questioned Document Examiners,* Vol. 9, No. 1, pp. 29-36.

Leung, S.C.; Cheng, Y.S.; Fung, H.T. and Poon, N. (1993). Forgery-Simulation. *Journal of Forensic Sciences*, Vol. 38**,** No. 2, pp. 402-412.

Masson, J.F. (1996). An Evaluation of Simulated and Traced Signatures and Consideration of the Potential for Determination of the Writer's Identity. Presented at the meeting of the American Academy of Forensic Sciences, Nashville, TN.

Osborn, A.S. (1929). *Questioned Documents,* 2nd edition. Chicago: Nelson Co.

Schuetzner, E.M. (2006). A Study of Attempted Signatures by Teenage Writers. *Journal of the American Society of Questioned Document Examiners, Vol.* 9, No. 2, pp. 95-101.

Srihari, S.N.; Cha, S.H.; Arora H. and Lee, S. (2002). Individuality of Handwriting. *Journal of Forensic Sciences*, Vol. 47, No. 4, pp. 856–872.