EasyChair Preprint

№ 8501

Numeric Approach in Handwriting Comparison

Marie Anne Nauer

July 20, 2022

Numeric Approach in Handwriting Comparison

Marie Anne Nauer
Institute for Handwriting Sciences
Scheuchzerstrasse 20, CH-8006 Zurich, Switzerland
m.a.nauer@bluewin.ch

Abstract. In forensic handwriting examination, the examinator is repeatedly confronted with very similar handwritings. In case of counterfeiting, such imitation is mostly executed as similarly as possible to the model. However, there are writers who by nature have a very similar style to their "model" or can put themselves in the person's shoes exceptionally well. In forensics, this often results in wrong judgements. This circumstance is to be countered by means of a more specific approach than usual. In consequence, in all these cases where highly similar handwritings occur we have to use comparison methods as e.g. systematic scales of quantified values on the basis of a highly differentiated quality analysis, followed by a systematic numerical approach. Some appropriate instruments have proven to be very useful and convenient in this context so as to achieve scientifically based assessments. In the following, some related methods will be presented for discussion.

1. Introduction

The need for a systematic registration and collection of data of various kinds is indispensable for scientific methods [4]. In the case of the recording of graphic facts in the daily practice, however, such a method is not yet particularly widespread. In the German-speaking and European countries, comparative handwriting examinations are usually conducted according to the *Best Practice Manual for the Forensic Examination of Handwriting* edited by the European Network of Forensic Science Institutes ENFSI [1]: It gives a detailed comprehensive description of best practice in this field. A large part of the manual is devoted to the instructions for the technical examination; however, it is often the case that these do not provide any further insights when we have to deal with handwriting comparison. Then it is often almost exclusively a matter of identifying graphic facts [3]. In addition to some investigations that can still be classified as technical, the manual also recommends identifying and comparing a number of 17 basic graphic features. However, there is no question of systematic or even numerically scaled recording.

Yet, for a systematic determination of graphic characteristics, it makes sense to first classify them according to specific criteria developed for a systematic handwriting analysis, as a basis for further numerical processing [2]. Practical examples will be used to present such numerical methods.

1.1 Literature review

Michel [6], one of the “fathers” of modern handwriting comparison in the German speaking area, in 1982 recommends a systematic approach in practice, but not any scaling of the findings obtained: he develops "a general system for recording script features", it must be "assumed that the basic components of script are sufficiently universal to cover the entire range of variation in manuscripts", and he then gives an overview of the classification of graphic features divided into nine basic components, which should represent both "delimitable functional units of the act of writing" and "the most practicable units of possible analysis for comparing the writing features", such as

1. line texture
2. pressure
3. flow of movement
4. movement guidance and shaping
5. direction of movement
6. vertical extension
7. horizontal extension
8. vertical surface structure
9. horizontal surface structure

Nezos [7] in 1994 lists 10 basic dimensions and subdivides them into numerous sub-dimensions; she further warns to pay special attention to a number of more than 30 individual elements that may indicate a forgery. There is neither any mention of scaling the findings.

Seibt [8, 9, 10] in the years of the 90s basically adopts Michel's classification, subdivides and defines the basic components in terms of subordinate features and illustrates them carefully. Furthermore, she lists 16 important aspects to be examined in order to assess the degree of congruence among two writings. A rating scale is not found here either.

Later [11] she mentions 7-point rating scales for 9 basic dimensions and 50 subdimensions, which are, however, only compared qualitatively, yet in form of very differentiated description, in the further course for comparative evidence. She keeps to emphasize on a strictly systematic proceeding [12, 13].

Viñals & Puente [15, 16] enumerate a large number of features and illustrate them in great detail, which is very helpful, especially for training; yet, a systematic registration is lacking as well.

The *Best Practice Manual for the Forensic Examination of Handwriting* improved its recommendation to compare 11 main characteristics in the version of 2018 to a number of 17 in the version of 2020 [1], yet, a recommendation for systematic or even numerically scaled recording and comparing is missing.

1.2 Summarizing conclusion of literature review

In summary, it can be said that a most detailed possible recording of graphic variables in the comparison of scriptures as well as a more or less explicitly mentioned systematic registration and comparison is mostly suggested to the expert by many authors, but the evidentiary collection of findings is carried out exclusively by means of subjective considerations. A continuous numerical approach has not been proposed so far and, thus, a consistent systematic assessment of the findings, in order to achieve an evidence-based evaluation as objective as possible, has not been thought through to the end.

2. Exposition of the problem - example of very similar handwritings (case 1): Pablo Picasso (1881-1973) and René Magritte (1898-1967)

As an example of a natural and unintentional similarity, the writings of Picasso and Magritte are cited.

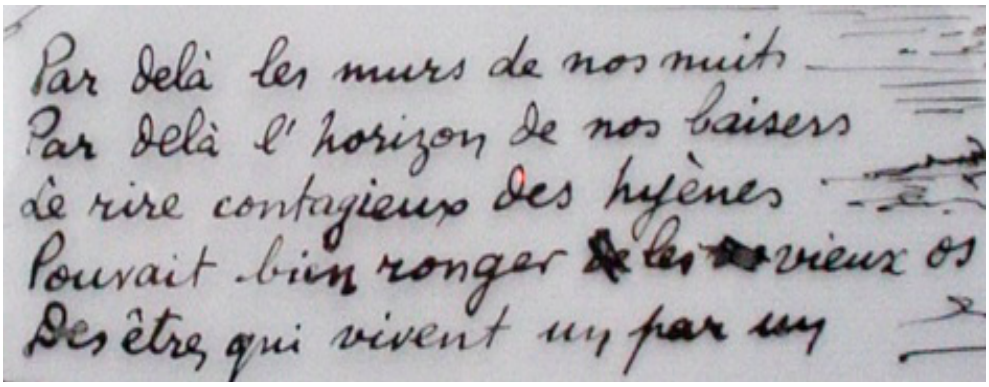

Figure 1. *Pablo Picasso: Letter excerpt in poem form (1936)*

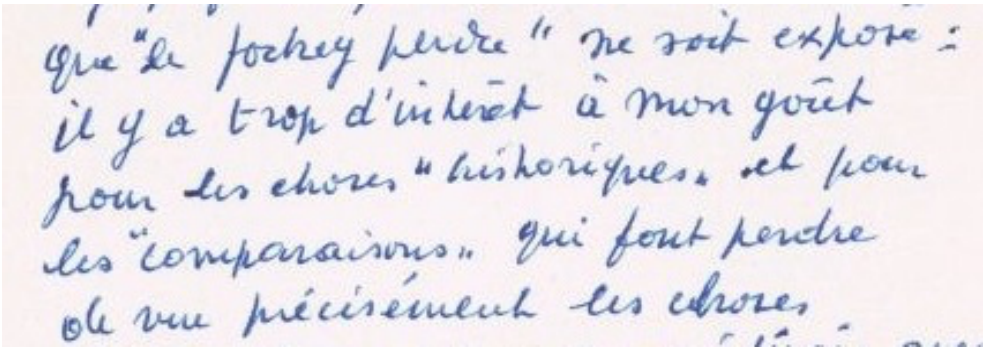

Figure 2. *René Magritte: Letter excerpt*

Although it is showed that completely identical writings do not exist [14, 17, 18], it is indeed not easy to distinguish between two writings that have so many similar forms.

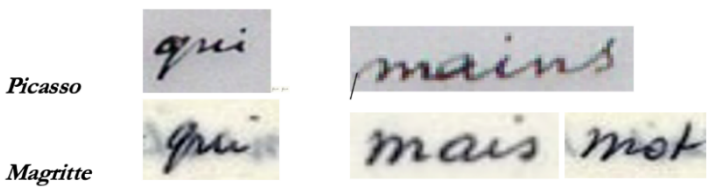

Figure 3. *Examples of very similar characteristics*

Here, for example, are highlighted

- In the word ***qui***:

Rightward hand connection in ***q*** /transition to ***u*** in a stiffened, arcadian form / actual form of the connection as an angle / next connection as a garland / downstroke of the ***i*** tends to be shorter / ***i***-dot set separately.

- In ***mains / mais, mot***:

Arcade shape of ***m*** / steeper position of the third downstroke in ***m*** / non-connection and shaping of the ***a***, the ***i*** and the ***s*** / raising of the last strokes of the ***n*** (***mains***) or the ***m*** (***mains, mot***).

Indeed, there are found 6 very similar formations with the word *qui* in only 3 letters, and 12 very similar formations with only 8 or 10 letters. If there would not be further possibilities for comparison available in this case, any expert would probably – and erroneously - conclude that it was the same authorship.

Looking at the whole letters, the impression is confirmed as far as both painters are used to mix text and drawings; yet, it might change somehow as also deviations are found. Anyway, the writings have to get analyzed systematically in detail.

3. Numerical-systematic examination methods

3.1 The numerical-systematic estimation of handwriting features

For the systematic determination all graphic features were first classified according to the five basic dimensions as: movement, form, space, pressure and stroke; most comprehensive as well as specific characteristics can be prevalently assigned to one of these dimensions. This compilation is not always compulsory, as individual characteristics often contain qualities of several basic dimensions, but it allows a better overview, as it is presented here. For the present differentiation chosen for practical purposes (which can of course be broken down further), there are between 3 and 30 in average, in this case between 6 and 21 characteristics per dimension, resulting on an average total of about 40 to 70 estimated items. The use of this number is useful and convenient in practice; however, Chernov [2] lists up to a thousand, strongly broken-down features to assess for scientific purposes.

Not all of them can be recorded in every manuscript, especially in signatures or short texts the number is limited. For numerical recording, each individual characteristic is classified on a seven-point scale[1]; subsequently, the congruence can be determined numerically from the assessed values.

Values can be compared simply arithmetically, and, by doing this, we obtain values of congruence or deviation: Hence, over all five dimensions with 42 scaled features, 9 positive and 47 negative congruence (deviation) points are resulting.

If these values of the present case are subjected to the X^2 test, the process obtains a significance value of slightly under 1 per cent, yet certainly under 5 per cent.

Table 1. *Five dimensions: congruence against deviation (non-weighted)*

Congruence Picasso/Matisse	Congruence C	Deviation D
Movement	0	11
Form	5	10
Space	3	16
Pressure	0	5
Stroke	1	5
Total	**9**	**47**

Significance of proportion ***C/D → p**~ 1%*** *(according to* X^2 *test, value: 6.446; critical value: ≥ 6.63)*
Significance of proportion ***C/D → p*≤ 5%*** *(according to* X^2 *test, value: 6.446; critical value: ≥ 3.84)*

3.2 The mitigating of results: deviation weighted

Yet, in order to avoid too strong deviations and to weaken the strong potential of a merely arithmetic deviation value, in many cases a deviation weighting is recommended. The definition may again be stronger or weaker, depending on the material available as well as on certain quality criteria.

Weighting definitions (proposal for use in practice):

The definition of ***congruence*** is equivalence of the values and is scaled by 1 congruence point.

The definition of the ***deviation*** is scaled as follows:

- 1 point from medium value → deviation value: 0
- 1 point on same polar side → deviation value: 1
- 2 points in general → deviation value: 1.5
- ≥3 points in general → deviation value: 2

[1] A seven-point scale might be defined from -3 to +3 for better visual retractability in daily practice, yet, it has to be transformed in a scale counting from 1 to 7 for further numeric and statistical calculation (1 meaning a very weak appearance, 4 a medium one, and 7 a very strong one).

This is particularly important in difficult cases:

- If the difference between the values of congruence and deviation remains significant, the result is even stronger.
- If the difference between the values of congruence and deviation is no more significant, the result may, but must not be weakened; anyway, it may just point to a strong tendency. This indicates that further qualitative investigations are compulsory and shows that the expert never is allowed to rely on one only instrument.

This process was done with the writings of Picasso and Magritte as well and the finding was that the mitigated version is no more significant yet pointing to a strong tendency.

Table 2. *Five dimensions: congruence against deviation (weighted)*

Congruence Picasso/Matisse	Congruence C_W	Deviation D_W
Movement	0	4.5
Form	5	5
Space	3	11
Pressure	0	3
Stroke	1	3.5
Total	**9**	**27**

Significance of proportion C_W / D_W → $p > 5\%$ *(according to* X^2 *test, value: 2.25; critical value: ≥ 3.84)*

3.3 A case with two possible authors (case 2)

In a case of a questioned testament, two persons were possible writers of the text. In order to find out the correct author, a numerical-systematic estimation of handwriting features was carried through.

Table 3. *Values of weighted congruence (C_W) & deviation (D_W) of three handwritings (X: questioned text; V: author 1; W: author 2)*

Characteristics	C_W X/V	D_W X/V	C_W X/W	D_W X/W	C_W V/W	D_W V/W
Total	**18**	**56**	**46**	**5**	**13**	**31**
Significance		**p*≤ 5%**	**p**≤1%**			p≥5%

Table 3 shows a significant deviation between the features of the questioned handwriting and the text of author V as well as highly significant congruence between X and author W. As a kind of background test, the features of the two possible authors were compared additionally; there is a strong tendency towards negative congruence, but no significant deviation. Yet, this finding shows that between the authors V and W there is a relatively high similarity and the investigation has a valid reason.

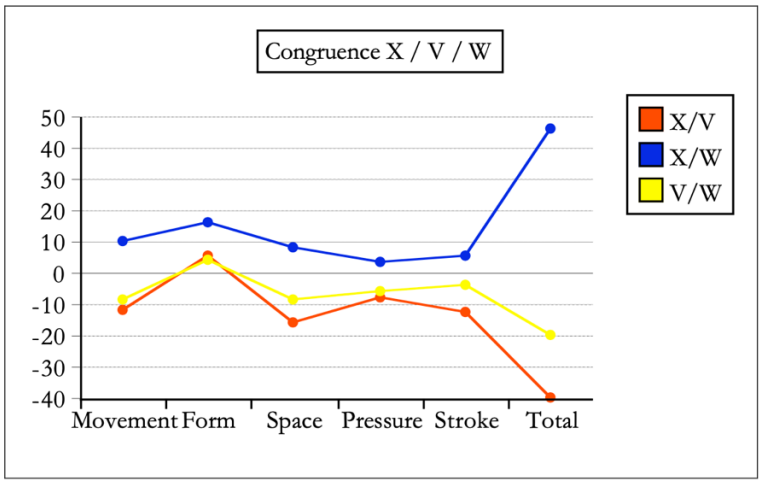

Figure 4. *Summarized congruence (positive and negative) for basic dimensions and in total of the scaled characteristics of three handwritings (X/V – bottom curve; X//W – top curve; V/W – medium curve)*

In Figure 4 the summarized weighted positive and negative congruence (deviation) shows a very clear picture leading to a very high probability that the questioned testament X was written by author W, and not by author V.

4. Cluster analysis

4.1. The cluster analysis

In certain cases, not only several manuscripts, but several groups of manuscripts are to be compared. This results in a large amount of numerical data that requires special processing.

A cluster analysis is used to uncover similarity structures between graph-theoretically classified groups that have not been pre-assigned and to re-identify them. In the process, the deviations from each other group are estimated on the basis of the graphological variables as described above. By the assigned values, the determined distance between the groups is given: The greater the distance, the further "away" the groups are from each other.

When dividing into clusters, the following applies: A formed cluster should be as homogeneous as possible, but at the same time differ as much as possible from the other clusters. In order to ensure this, certain prerequisites regarding the data basis should be taken into account, such as missing values in the data set can falsify the analysis. Therefore, they must always be as complete as possible, and thus, in certain cases, not all assessable or measurable characteristics can be included but just those that are found in every group.

Since in practical applications the objects to be analyzed are already very similar to each other, it is best to work with a so-called hierarchical-divisive method, i.e. each variable is mapped as a separate cluster. These are sorted in a hierarchical sequence, and the two most similar clusters are successively combined into a new cluster until all objects are in one category. In this way, a kind of tree results whose branches are closer or further apart, depending on the distance determined.

The measure used is defined by the applied scale. The basis however is again the registration by a seven-point scale as before.

4.2. Exemplifying case 2

In the case of the questioned document in case 2, there were some additional documents Y that could not be, due to the relatively high similarity between the two possible authors, clearly assigned to one of them. The cluster analysis (Figure 5) yet shows a clear distance of V to all the other writings building one big group. This result reinforces the findings of the systematic analysis confirming that V cannot be the author of the questioned documents, and classifies the writings Y among the group of X and W.

The picture shows a tree with several branches, with the single branch “V” lying distinctly apart from the all the others branches which form one group with related branches among each other.

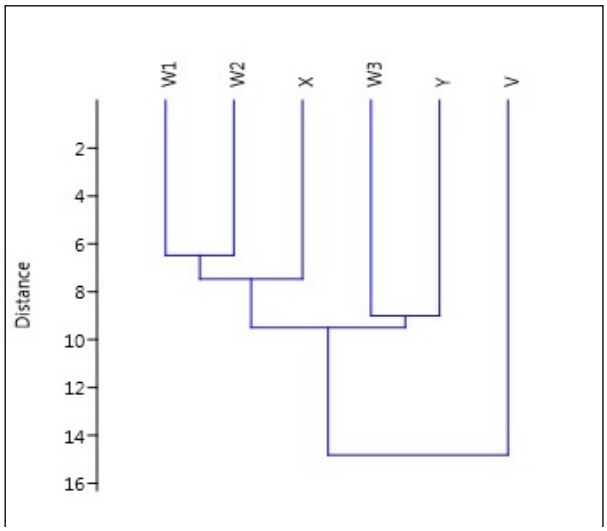

Figure 5. *Cluster analysis for the writings X, V, W1, W2, W3, Y*

4.3. Example with different groups of writings (case 3)

In a handwriting comparison case, there were five distinguishable groups of comparative manuscripts available: two postcards with spontaneous cursive writing (V1, V2), with dates preceding the questioned document; furthermore, the suspected person was asked by the police to produce cursive handwriting samples, which were written particularly carefully (V3-V5). In the further course he prepared a draft in printed writing (V6), he described the events by several handwritten cursive letters to the expert (V7-V9) and submitted a note with a short spontaneous note in printed writing (V10).

Since all these groups of manuscripts, in a certain way, appeared different from each other, they were divided up as described and each group was systematically classified according to 43 characteristics on a seven-point scale.

The table 4 shows the exemplary classification of 8 characteristics associated to movement as shown in the table, scaled for 6 groups of handwritings; the results leading to a table with 83 values. Accordingly, the four other dimensions were classified.

Table 4. *Case 3 - classification of signs associated to movement*

Example Handwriting comparison	X1	V1, V2	V3-V5	V6	V7-V9	V10
Signs associated to movement						
Movement accentuation	5	5	4	4	4	5
Strength of impetus	5	6	4	5	5	5
Homogeneity of flow	2	6	7	6	7	5
Looseness vs. bond	2	5	6	6	6	5
Tempo: Speed of Stroke	6	5	5	5	5	5
Tempo: Speed of Success	5	6	5	6	6	6
Connectedness: degree	3	3	6	1	6	1
Leftward/rightward trend: general	5	6	5	5	5	5

The systematic classification of all 43 characteristics leads to a table of 258 values. In this case, a fairly clear picture emerges: the questioned handwriting X shows a fairly large distance to all the other writings (Figure 6).

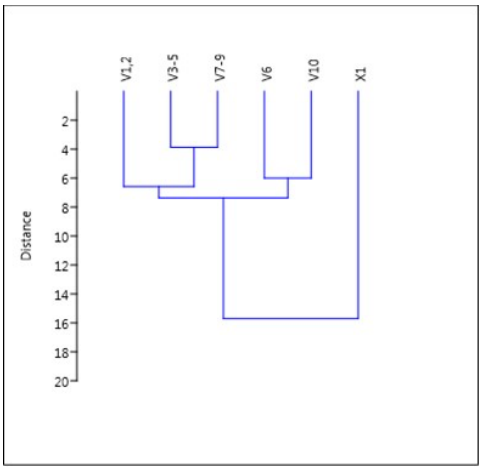

Figure 6. Case 3 - cluster analysis

The comparative manuscripts, however, even if by only small distance to each other, divide again into two groups: the cursive manuscripts and the ones with printed letters (V6, V10). This corresponds very well to the other findings in this case and shows a clear deviation of the characteristics of the questioned handwriting X.

4.4. A historical question (case 4)

A historian sends a number of signatures for comparison: eight of them date from the years 1911-1924 and are in the name of "C"; ten signatures date from the years 1946-1965 and are in the name of "H". No material is available from the period between 1924 and 1946. The question is whether the two groups were written by one and the same person or by two different persons. According to historical tradition, it should be the same person.

Obviously, however, researchers doubt the correspondence of the two groups of writings; conspiracy theories are apparently also circulating, pointing to a disappearance of the first person "C" and a mere pretense of his reappearance in the form of a double. The whole thing is complicated by the fact that the person in question demonstrably used a large number of different names throughout his life. In the years between 1924 and 1946, the person in question spent a long time abroad; however, it is not known why there are no signatures from this long period.

Technical problems:

- The material is available only in highly deficient quality (downloads from the internet in partly poor resolution).
- There are almost only signatures and only a few writings available; the examination must therefore be limited to checking the signatures.
- The origin and authenticity of the documents cannot be verified by the expert.

An initial examination of the signatures shows that their design varies considerably; the chronological list shows an uneven distribution of different signature patterns either.

In order to gain an initial overview, in a first step 33 individual characteristics are examined for each signature, weighted and classified in a scale of 7 points; on the basis of this material, a cluster analysis could be carried out. In this example, we are dealing with numerical material from 18 signature protocols with 33 characteristics each, thus amounting in a number of 594 items.

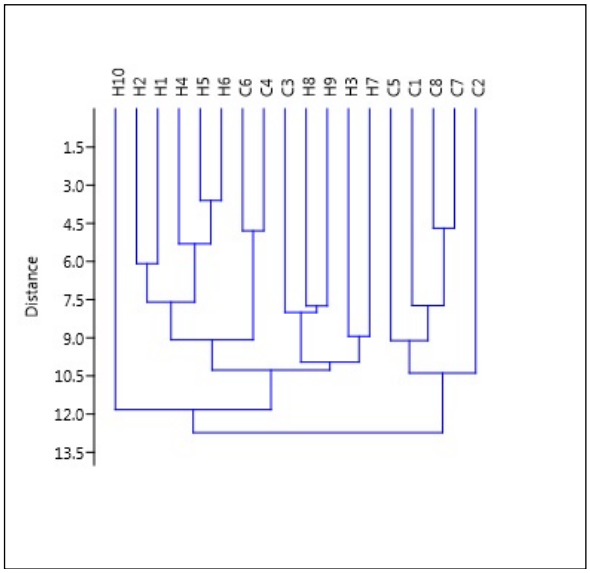

Figure 7. *Cluster analysis for case 4*

The picture of the resulting cluster analysis shows two main groups with regard to the distribution of the signatures or their similarity to each other, whereby within one of them there are only found writings of group "C" (C1, C2, C5, C7, C8, on the right) while the other one (on the left) includes two main groups, one of which in turn contains only one writing "H" (H10, on the far left) as well as a mixed group.

The isolated position of H10 can be explained by the fact that it is the last in the chronological order, with a ten-year distance from the next, and shows clear signs of age-related change.

The mixed group consists of three scripts of group "C" (C3, C4, C6) and the remaining nine scripts "H" (H1-H9). This mixture cannot be further explained even by re-examining the signatures individually.

Thus, the result is unclear and no tendencies can be described, let alone conclusions drawn. It therefore makes no sense here to carry out even more elaborate analyses without qualitatively and quantitatively good material, and without further information on the origin of the signatures. The conclusion is therefore: *Non liquet* - unfortunately to the great disappointment of the historian.

5. General Discussion

In the comparison of very similar handwritings it is often difficult to obtain a clear result. Of course, a certainty of 100% is never possible *a priori*, the method always remains in the domain of probability yet to assess as accurate as possible [5]. Nevertheless, there are not found authors who suggest a consequent systematic *qua* scientific approach for graphic features even if they do for physical, chemical, and other similar investigation in handwriting comparison. Yet, a systematic numeric registration as well as its mathematic working up allows a better approach to a higher degree of probability.

The numeric tools presented and proposed here should not be considered as solely conclusive evidence in a forensic analysis any more than any other method. Yet, experts do have to apply and to develop as many instruments as possible. Hence, as an extension of the *repertorium* of methods, these instruments can support the evidence in certain cases.

References

1. *Best Practice Manual for the Forensic Examination of Handwriting* (2020). European Network of Forensic Science Institutes. ENFSI-BPM-FHX-01. (Ed. 03).
2. Chernov, Y. (2018). Formal Validation of Handwriting Analysis. *In: Y. Chernov, M. A. Nauer (Eds.). Handwriting Research – Validation & Quality* (pp. 37-69). IHS Books.
3. Edelmann, J (1897). *Die Schrift-Expertise im Strafprozess: Ein Mahnruf gegen die Verurteilung Unschuldiger. Mit zwei Beispielen*. Zürich.
4. Horst, D., & Nauer, M. A. (2019). *Graphologische Gutachten in Theorie und Praxis aus arbeitsrechtlicher Sicht.* DIKE.
5. Köller, N., Nissen, K., Riess, M., & Sadorf, E. (2004). *Probabilistische Schlussfolgerungen in Schriftgutachten. Zur Begründung und Vereinheitlichung von Wahrscheinlichkeitsaussagen im Sachverständigengutachten* [Probabilistic conclusions in written expert reports. On the justification and standardisation of probability statements in expert reports]. BKA.
6. Michel, L. (1982). *Gerichtliche Schriftvergleichung. Eine Einführung in Grundlagen, Methoden und Praxis*. Walter de Gruyter & Co.
7. Nezos, R. (1994). *Judicial Graphology*. Scriptor.
8. Seibt, A.(1992). Schriftpsychologie und Schriftvergleichung. *Zeitschrift für Schriftpsychologie und Schriftvergleichung, 56*, 2, 92-105.
9. Seibt, A. (1997). Methode schriftvergleichender Untersuchungen. *Zeitschrift für Schriftpsychologie und Schriftvergleichung*, *61*, 87-101.
10. Seibt, A. (1999). *Forensische Schriftgutachten*. Beck.
11. Seibt, A. (2000). Schriftvergleichende Befunderhebung: Skalen und Messtechniken. *Zeitschrift für Schriftpsychologie und Schriftvergleichung, 64*, 38-53.
12. Seibt, A. (2003). Berührungspunkte und Unterschiede zwischen forensischer Handschriftvergleichung und Schriftpsychologie. *Zeitschrift für Schriftpsychologie und Schriftvergleichung, 67*, 1, 2-23.
13. Seibt, A. (2004). Qualitätsmerkmale forensischer Schriftgutachten. *Zeitschrift für Schriftpsychologie und Schriftvergleichung 68,* 1, 44-62.
14. Srihari, S. N., Cha S.-H., Arora, H., & Lee, S. (2002). Individuality of handwriting. *Journal of Forensic Sciences*, 47,4, 1–17.
15. Viñals, Francisco & Puente, Maria Luz. Pericia caligráfica judicial. Barcelona 2001
16. Viñals, Francisco & Puente, Maria Luz. Grafología criminal. Barcelona 2009
17. Zhang, B., & Srihari, S. N. (2003). Binary vector dissimilarities for handwriting identification. In *Proceedings SPIE, Document Recognition and Retrieval IX*, pp. 155–166.
18. Zhang, B., Srihari, S. N., & Lee, S. (2003, August 3-6*). Individuality of Handwritten Characters.* 7th International Conference on Document Analysis and Recognition, Edinburgh, Scotland.