J18 A Social Science Paradigm for Forensic Handwriting Identification

Carole E. Chaski, PhD, ALIAS Technology, LLC, Georgetown, DE 19947*

Learning Overview: After attending this presentation, attendees will be able to evaluate the possibility of applying a social science paradigm to forensic handwriting identification using the model presented.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by providing: (1) principles for applying the social science research paradigm to forensic handwriting identification and other pattern-recognition forensic sciences, and (2) ways that this paradigm could shift current practices.

Forensic handwriting identification is practiced within governmental crime laboratories and independent consultancies and was the first forensic science to come under the scrutiny of the *Daubert* factors for scientific evidence.[1] In the past 20 years, there have been *Daubert* challenges, generally with mixed results: some exclusions, some inclusions, and some restricted inclusions of testimony.[2,3] The forensic document examination community, the legal community, and the academic research community have offered multiple solutions for meeting the *Daubert* factors, including the development of software for fully automated handwriting identification, such as CEDAR-FOX and FLASH-ID, proficiency experiments, and consensus-based standards in groups such as Technical (Scientific) Working Group on Questioned Documents and the Organization of Scientific Area Committees subcommittee for forensic document examination.[4,5] All of these solutions stem from the desire to provide forensic document examination, and especially handwriting identification, with an acceptable and feasible scientific foundation, objectivity, and mitigation of confirmation bias.

This presentation focuses on a social science research paradigm that can be applied to handwriting identification as well as other pattern recognition techniques. This social science paradigm is already accepted within major social sciences (e.g., psychology, linguistics, sociology).[6-9] While social sciences (e.g., psychology, linguistics, sociology) are known for strong quantitative analysis and laboratory experimentation, they also provide research guidelines for qualitative assessment based on educated judgments. The main purpose of this paradigm within the social sciences is to produce objective results when the analysis is based on qualitative assessment. This paradigm involves an administrator and a small team of at least three experts. The experts work independently ("blindly") using a checklist as well as unstructured notes for the analysis of items, performing a qualitative assessment of the data. This paradigm provides a measure of reliability for the assessment using well-known and standard statistical procedures for inter-rater reliability (i.e., agreement among qualitative assessments).[10] The paradigm's experimental design or workflow can easily be replicated. Finally, this paradigm can be adopted in both governmental laboratories and independent consultancies, which could lead to standardization in the field.

The social science paradigm has already been employed in forensic linguistics for the qualitative analysis of suicide notes and threat letters.[11] The workflow for this paradigm in forensic linguistics has been implemented in software to insure that: (1) the administrator manages the team objectively, (2) the team members actually do work independently ("blindly") and cannot be influenced by each other's assessment (thus avoiding group-induced confirmation bias), and (3) the administrator runs the automated quantitative assessment of the team's collective decision and inter-rater reliability. The report based on this paradigm includes both the collective decision and the inter-rater reliability so that the strength of the decision can be evaluated using statistical information.

This social science paradigm differs in several important ways from current practices in both governmental crime laboratories and independent consultancies. Thus, there may be cultural resistance to adopting this paradigm. On the other hand, if adopted by the forensic document community, forensic handwriting identification could align with standard social sciences.

Reference(s):

1. *United States v. Starzecpyzel*, 880 F. Supp. 1027 (S.D.N.Y 1995).
2. Michael D. Risinger. Cases Involving the Reliability of Handwriting Identification Expertise Since the Decision in *Daubert*, 43 *Tulsa L. Rev*. 477 (2013). Available at: http://digitalcommons.law.utulsa.edu/tlr/vol43/iss2/11.
3. Zlotnick J., Lin J.R. Handwriting Evidence in Federal Courts—From *Frye* to *Kumho*. *Forensic Sci Rev. 2001*;13:87–99.
4. Kam M., Wetstein J., Conn R. Proficiency of Professional Document Examiners in Writer Identification. *J Forensic Sci.* 1994;39:5–14.
5. Sita, J, Found, B, Rogers, D. K. Forensic Handwriting Examiners' expertise for signature comparison. *J Forensic Sci*, Sept. 2002, Vol. 47, No. 5:1117-24
6. Berg, B.L. 1988. *Qualitative Research Methods for the Social Sciences.* New York: Pearson.
7. Armstrong, D., Gosling, A., Weinman, J., and Mateau, T. The Place of Inter-Rater Reliability in Qualitative Research: An Empirical Study." *Sociology*. 31(3) 597-606.
8. Milroy, L. 1987. *Observing & Analyzing Natural Language.* New York: Blackwell.
9. Richardson, J.T.E. (ed). 1996. *Handbook of Qualitative Research Methods for Psychology and the Social Sciences.* New York: Wiley-Blackwell.
10. Hallgren, K.A. 2012. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol.* 8(1): 23-34. Available online at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3402032/pdf/nihms372951.pdf.
11. Chaski, C.E. and Huddle, D. Is This a Real Suicide Note? Authentication Using Statistical Classifiers and Computational Linguistics. *Proceedings of the American Academy of Forensic Sciences*, 63rd Annual Scientific Meeting, Chicago, IL. 2011.

Forensic Handwriting Identification, Social Science, Inter-Rater Reliability