J24 The Interaction of Writing Profiles and Automated Scoring Rules

Cami Fuglsby, MS, South Dakota State University, Brookings, SD 57007; Michael Caligiuri, PhD, University of California, San Diego, CA 92093; Danica Ommen, PhD, Iowa State University-Statistics, Ames, IA 50011; Christopher P. Saunders, PhD, Brookings, SD 57006; JoAnn Buscaglia, PhD, Federal Bureau of Investigation Laboratory, Counterterrorism and Forensic Science Research, Quantico, VA 22135*

Learning Overview: After attending this presentation, attendees will better understand the importance of statistical invariance as it relates to the development of automated scoring rules for handwriting verification systems. The Random Non-Match Probability (RNMP) is the chance that two writing samples written by the same writer are declared a non-match by a given biometric scoring method. Some important questions that arise are whether or not this chance is the same for every writer in a population of writers and whether or not it is affected by the number of words in the writing samples.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by delving into the design and validation of automated handwriting identification systems to writing samples, the statistical methods, and interpretations based on scores.

Hypothesis Statement: Designed studies of handwriting will allow researchers to identify scoring rules that have an RNMP that varies as a function of a writing profile.

In the construction of Automated Handwriting Identification Systems (AHIS), pairwise scoring rules play an important role associated with characterizing the performance metrics of the system. Most AHIS for forensic source identification are optimized to recommend the order in which an examiner should search a candidate list of known writers to identify the source of a writing sample with an unknown writer. In effect, they recommend the source of a writing sample that an examiner will then verify. A natural measure of the performance of a scoring rule is the associated Random Match Probability (RMP). The RMP is the chance that two writing samples from different writers are declared a match—this is analogous to the "likelihood of a chance match" in the logical approach to handwriting analysis. Unfortunately, most handwriting AHIS are not optimized for minimizing the RMP and are instead focused on minimizing the time to find a writer in database of writers. This means that, if one wishes to compare two writing samples in a head-to-head comparison to measure the similarity with a given AHIS, then the researchers will usually need to design a strategy for encoding the output of the system in order to develop an interpretable scoring rule (by interpretable, this study means that a score has meaning that is consistent across writing profiles, or that the RNMP for two writing samples written by writer A is the same as two writing samples written by writer B). This work will discuss various strategies for constructing pairwise dissimilarity scores and characterizing the performance of the resulting scoring rules.

Using writing samples collected under a National Institute of Justice (NIJ) -funded research project, this study explores the properties of various pairwise scoring rules for the results of the FLASH ID system. This study also focuses on the scores' dependence on the amount of handwriting (in each writing sample) that is made available to the system. Another important focus of this research is on the visual display of the statistical methods to convey the performance of the scores.

Source Identification, Handwriting, Biometric Performance Metrics