****

Vlad Atanasiu, Expert Bytes: Computer Expertise in Forensic Documents – Players, Needs, Resources and Pitfalls, Boca Raton, FL: CRC Press, 2013

****

> ADDENDA & ERRATA

Addenda

Matley 2008 reference [p. vii] — Marcel B. Matley points out the following regarding the reference to one of his articles: "You cite my review of Shrihari's paper on handwriting of twins at the beginning. Unfortunately, you must have had an improper copy since what you attribute to me is not what my critical review says. I do not deny computers can be of value, I only demonstrate how his particular study shows that his specific program and research is much mistaken and off base by a long ways." [email to the author, 2013.10.28]

What I found interesting in Matley's point of view is that it can be seen as an epitome perspective on computational expertise of forensic documents (FD), beyond characterizing a specific article, the one by Srihari. As Matley made it clear, it is not computers *per se* that he is battling, but "misbegotten computerized research." He writes: "After 20 or 30 years of tax and other funds being dumped into the bottomless pit of misbegotten computerized research, we have nothing of practical use to show for it." [same reference]

Computational paleography reference [p. 3n1] — A review of the current state of issues and perspectives in computational paleography is offered in Tal et al. 2013. — See new references at the end of the addenda.

On the writing style of law scholars [p. 5] — Of which it was famously said: "I am the last one to suppose that a piece about the law could be made to read like a juicy sex novel or a detective story, but I can not see why it has to resemble a cross between a nineteenth century sermon and a treatise on higher education." [Rodell 1936: 41]

User-centered design [p. 30] — Kansei engineering, a Japanese design philosophy aiming to elicit positive affects from users in respect to products, is particularly suited to issues of user-centered design.

Know-how + Tools [p. 45] — In complement of technical references, a recommended reading is NAS 2010 for its critical perspective on biometrics, the close relative of FD, which is envisioned as a system engineering problem, one that is not infallible, and is part of a social, cultural, and legal context.

Sampling [p. 47, replaces the second paragraph] — *Sampling* — Let's not forget that the choice of sampling criteria depends on what one wants to do with the samples once collected. A nation-wide sampling cast like an opinion poll utilizing census criteria (age, gender, education, income, and so forth) is an appropriate model for a software employed at country level (the United States

in one study), but less so for a local usage if the demographics differ from the national average (and apart from the fact that we do not know what the sample size appropriate for studying handwriting should be).[1] Indeed, numerous FD questions require specialized datasets, which can shed light on the individuation power of Latin script compared to Arabic or Chinese, distinguish between the standardized writing styles of calligraphers or monitor that of people suffering from Alzheimer's, penetrate the intriguing peculiarities of Basque terrorist handwriting[2] or the impact tactile screens on the ability to distinguish signatures (notoriously bad for anyone having had to sign a FedEx receipt).[3] Research of the past 15 years has also shown the importance of considering the production conditions when assessing writing individuality. For example, its value is lower for simulated writing collections when compared to purely genuine samples[4] or twins' handwriting compared to that of nontwins', an astute experiment since long favored by psychologists and FDEs.[5]

Given the thorny issue of sampling, researchers might want to look for alternatives. A class of such methods, called non-probabilistic because they don't rely on random sampling, are known from the area of survey design. One of them in particular, "purposeful sampling," is interesting for the topic of handwriting.[6] Contrary to probabilistic sampling, purposeful sampling is a qualitative approach that seeks to understand the mechanism behind a particular distribution, in order to provide explanations instead of merely quantitative descriptions. The model mechanism can then be used to simulate distributions, thus bypassing sampling (in the right conditions and to a certain degree, since even a model needs to be grounded in data). Single case studies are another option, with a long history in medicine, where FDEs can look for methodological inspiration when confronted with such questions as documenting the transformation of an individual's handwriting in time.[7]

1. Srihari et al. 2002: 858–859 propose to adopt the size of opinion polls, around 1,100 subjects, but Sacks 2003a: 916–917 points to medical experiments commonly with as few as 10–20 patients and epidemiological studies carried out on tens of thousands of individuals; NAS 2010: 30–31 discusses the global world scale. — **2.** Bisotti 2011. — **3.** Harralson 2013: 62–64. — **4.** E.g., Found and Rogers 2007. — **5.** References in Harrison and Seiger 2003, Huber and Headrick 1999: 196–198, study in, e.g., Srihari et al. 2008, critique in Page et al. 2011: 15; compare with sampling issues in the FBI mtDNA database in Kaestle et al. 2006. — **6.** On purposeful sampling see Leyden 2013: 113–128, Patton 2002: 45–46, 230–247; on the long history of sample vs model based inference see Brewer and Gregoire 2009; a practical introduction to sampling is Henry 1990; for reference works see Pfeffermann and Rao 2009, Gideon 2012. Usually sampling methods are quantitative, produce descriptions, are used for measuring prevalence of phenomena, and allow generalization of results to larger populations; model methods are qualitative, explanatory, generate mechanisms, and few of them are statistically representative. In terms of terminology the approaches are variously called in statistics and the social sciences random, probabilistic, nonparametric, design based, as opposed to nonrandom, nonprobabilistic, parametric, model or prediction based, to which the medical literature adds the terms homeothetic and extensive (general and group), respectively ideographic and intensive (individual). — **7.** Although concerned with clinical psychology, scientists from many other fields can benefit from the excellent overview of the history, issues, and methods of single case experiments compared to group and randomized studies is Barlow et al. 2009: 1–29 [history], 31–59 [issues], 46–52 [possibility of generalization from single cases].

Descriptors [p. 48] — Psychophysical experiments represent a possible way forward to understand what features are used by the human visual system for comparing script styles. [Features that allow the distinction between characters, an important aspect for understanding reading mechanisms, have been studied in psychology (references in Sanocki 2012: 135–136).]

Handwriting individuality [p. 55–57] — The US National Academy of Sciences calls for the development of a "science of human individual distinctiveness" and sees it as "essential to effective and appropriate use of biometric recognition," given the complexity and social impact of such systems. [NAS 2010: 4] On the dogma of biometric recognition and anthropometry see NAS 2010: 23–52, respectively Gould 1996.

Dataset representativeness [p. 61] — That "population sampling issues that affect performance in practice may not be fully appreciated," has also been highlighted by the US National Academy of Sciences and contrasted with the "very sophisticated statistical methods are used for the signal analysis and pattern recognition aspects of biometric technology." [NAS 2010: 17]

DNA profiling [p. 72n4] — See Zimmer 2013 on multiple genomes (mosaicism and chimerism).

On interdisciplinarity [p. 78n2] — In fact, the issue of interdisciplinarity arose from the very inception of computing. In the words of Grace Hopper, in the 1940s one of the first programmers and central figure to the invention of software compilers and the COBOL programming language, "We had to learn their vocabularies in order to be able to run their problems. I learned languages of oceanography, of this whole business of minesweeping, of detonators, of proximity fuses, of biomedical stuff." [Beyer 2009: 55]

HG4FDE, Anonymous [p. 124] — The software is now called "Masquerade", and has its website at http://www.nitesrl.com/products/masquerade.

References

Barlow David H., Matthew K. Nock, Michel Hersen (2009) — *Single Case Experimental Design: Strategies for Studying Behavior Change*, Boston, MA: Pearson Education.

Beyer Kurt W. (2009) — *Grace Hopper and the Invention of the Information Age*, Cambridge, MA: The MIT Press.

Brewer Ken, Timothy G. Gregoire (2009) — *Introduction to Survey Sampling*, in Pfeffermann and Rao 2009 (29A): 9–38.

Gideon Lior, ed. (1996) — *Handbook of Survey Methodology for the Social Sciences*, New York: Springer.

Gould Stephen Jay (1996) — *The Mismeasure of Man*, New York: W.W. Norton.

Henry Gary T. (1990) — *Practical Sampling*, Newbury Park, CA: Sage.

Layder Derek (2013) — *Doing Excellent Small-Scale Research*, Thousand Oaks, CA: Sage.

Nagamachi Mitsuo (2011) — *Kansei/Affective Engineering*, Boca Raton, FL: CRC Press.

Nagamachi Mitsuo, Anitawati Mohd Lokman (2011) — *Innovations of Kansei Engineering*, Boca Raton, FL: CRC Press.

NAS (National Academy of Sciences, Whither Biometrics Committee) (2010) — *Biometric Recognition: Challenges and Opportunities*, Washington, DC: The National Academies Press.

PATTON Michael Quinn (2002) — *Qualitative Research & Evaluation Methods*, Thousand Oaks, CA: Sage, 3rd ed.

PFEFFERMANN Danny, C. R. Rao, eds. (2009) — *Sample Surveys: Design, Methods and Applications* (vol 29A), *Inference and Analysis* (vol 29B), Amsterdam: North-Holland.

RODELL Fred (1936) — "Goodbye to Law Reviews," *Virginia Law Review*, 23: 38–45.

SANOCKI Thomas, Mary C. Dyson (2012) — "Letter Processing and Font Information During Reading: Beyond Distinctiveness, Where Vision Meets Design," *Attention, Perception, & Psychophysics*, 74: 132–145.

SCHÜTTE Simon (2005) — *Engineering Emotional Values in Product Design: Kansei Engineering in Development*, PhD thesis, Linköping: Department of Mechanical Engineering, Linköping University.

TAL Hassner, Malte Rehbein, Peter A. Stokes, Lior Wolf (2013) — "Computation and Palaeography: Potentials and Limits (Dagstuhl Perspectives Workshop 12382)," *Dagstuhl Manifestos*, 2(1): 14–35.

ZIMMER Carl (2013) — "DNA Double Take," *The New York Times*, online, accessed 2013.10.03, http://www.nytimes.com/2013/09/17/science/dna-double-take.html.

Errata

P. 99 — For “2. forgery ○ (class)“ read “2. forgery ● (class)“.