

APPROACHES TO EVALUATION AND REPORTING OF EXPERT EVIDENCE PANEL DISCUSSION:
“The logical approach to evidence evaluation and reporting”

R. Brent Ostrum

This panel discussion will present a brief overview of the frequency of occurrence and likelihood ratio approaches to evaluating (handwriting) evidence, and their respective advantages and limitations. Reporting of evidence using these approaches will also be discussed, as well as potential future conclusion terminology. Time will be set aside for questions, comments and views from the audience.

Questions for panel participants to address

- Briefly describe the evidence evaluation/reporting approach you are speaking on
- What scientific literature/research is there to support this approach
- What are its advantages/strengths
- What are its disadvantages/limitations
- Who currently uses this approach
- What barriers do you see to its implementation

[DISCLAIMER: The positions and opinions of ease expressed in this document are those of the author. Unless otherwise stated, they do NOT represent or reflect the official policy of the CBSA Laboratory, the FDE section, the Agency (CBSA), the Government of Canada or any other agency or organization with which I am affiliated or associated.]

Brief description:

The “logical approach to evaluating handwriting evidence” is the focus of my presentation. I prefer the term “logical approach”, rather than the ‘likelihood-ratio’ or ‘LR’ approach, because a formal or strict LR approach requires numeric data – something we do not presently have and which presents significant challenges for the future.

From my perspective the logical approach works perfectly well when viewed primarily as a system of logical reasoning that can be expressed or explained using a verbal construct that is roughly equivalent to the LR concept. Basically, the opinion expressed at the end of the process will be a statement that relates to the probability of the evidence given competing propositions; rather than a statement about the propositions themselves. The precise wording of that opinion may vary as will be explained later.

The logical approach is more than just a new or different type of conclusion wording. Rather, it is a more complete approach to the entire evaluation process; one that applies equally to any type of evidence or information. ‘Conclusions’ that result from this evaluation process are straight-forward and are a direct reflection of the process itself.

A key objective for any forensic evaluation process (and any conclusion resulting from it) is that it must be logically coherent and sound. More specifically, there are four **essential** requirements that

should be met by any evaluation and reporting scheme – these are Balance¹, Logic², Robustness³ and Transparency⁴.

It is important to understand that this approach is not a panacea and will not address all of the issues and criticism we face as forensic practitioners. It can, however, help when dealing with our critics and some of their arguments. It can also help to address some human factors issues relating to bias, context management and interpretation.

Key points:

- Every expert opinion is based on an individual’s knowledge and training, experience, etc. which informs their beliefs about the findings and observations (alternatively, the “evidence” observed in an examination)
- Every expert opinion is based on uncertain information and, hence, is also uncertain to some degree. Uncertainty must be addressed through the application of probability and logic. Our reasoning/scientific belief can be expressed, either implicitly or explicitly, using probabilistic terms that are either quantitative or qualitative in nature.
- The basic rules for any evaluation are simple:
 1. Evaluation always occurs within a framework of information (that is, it is contextualized),
 2. There must be at least two competing propositions (ideally, representing the positions being argued by the parties), and
 3. The expert evaluates the evidence given the propositions, and not the propositions directly, and expresses their opinion accordingly.
- In mathematical terms, the (odds) form of the ‘construct’ is: $LR = \frac{p(E | H_1, I)}{p(E | H_2, I)}$

Where E = evidence/findings, I = framework information, and H_x = competing propositions

¹ ‘Balance’ means that the evidence/findings should be evaluated given at least one pair of competing propositions; ideally with the first proposition based upon one party’s account of the events and the latter based upon an alternative account.

² ‘Logic’ means that the evaluation process must be one that speaks first to the probability of the evidence/findings given the propositions (plus relevant background information), and not the probability of the propositions given the evidence/findings (plus background information). This is essential to ensure there is no inappropriate or unjustified transposition of the conditional since proper or correct transposition of the conditional requires information generally not within the scope of the examiner.

³ ‘Robustness’ means simply that the evaluation process must be capable of sustaining scrutiny or review by other experts through review or cross-examination. It should be based upon sound knowledge and experience of the evidence type including the use, when available, of pertinent databases, published data or ad hoc case based experimentation. In other words, ‘robustness’ refers to the scientist’s ability to explain the grounds for their opinion based upon their degree of understanding of the particular trace type and its probability of occurrence in the relevant ‘population’ relating to each of the competing propositions.

⁴ ‘Transparency’ applies to all facets of the examination and evaluation and means that the entire process should be demonstrable and recorded so as to permit proper review and assessment. Worknotes should clarify all relevant aspects of the evidence including the interpretation and evaluation of that evidence in terms of the competing propositions. The report should be written in way that is suitable for a varied audience (i.e. participants in the justice system).

- There are many ways that this concept (or its non-mathematical equivalent) could be expressed or conveyed verbally. The ENFSI guide provides some examples and others can be found in the literature.
- At the present time, and in the absence of research to direct otherwise, a structure that speaks to the “relative degree of support provided by the evidence” seems to have the broadest appeal amongst practitioners who advocate for this approach. In other words, the expert evaluates and expresses their belief in terms of the degree of support provided by the evidence for one proposition over another competing proposition. Generally, the degree of support is described using some standard set of modifiers.
- An example set **might** be the following (wording from strongest to weakest degree of relative support):

The evidence provides very strong support for proposition X over proposition Y.
The evidence provides strong support for proposition X over proposition Y.
The evidence provides moderate support for proposition X over proposition Y.
The evidence provides more support for proposition X than for proposition Y and the level of that support, while stronger for X than for Y, is limited/weak.
The evidence provides approximately equal support for proposition X and proposition Y.

- A ‘scale’ such as this could be adopted now and it would address many of the logical issues and problems that exist with our existing terminology. However, it is not perfect by any means. It would only be a stop-gap solution since research is still required to see what **specific numbers of levels** and what **specific wording** would be optimal for the purpose of correct and effective communication. Work by Martire (2014) and others has already shown that subtle word choices (or changes in semantic structure) can have dramatic effects on the meaning and understanding of the conclusions. I should note that this applies equally or moreso to our existing terminology so this is NOT an issue restricted to the logical approach.

Scientific literature/research to support the approach:

The basic analyses and examination processes an examiner uses to assess questions of authorship remains almost the same when applying the logical approach. As a result, most of our existing literature in support of claims of expertise will remain valid and applicable. That does not change simply by adopting this approach for the overall evaluation process.

Of course, validation is one area where our discipline can use more work. In the context of the present discussion validation could be considered in terms of at least two discrete elements:

1. The skills expressed and claimed by FDEs – each of our claims suggests the application of one (or more) methods or techniques. To some degree, each of those methods requires validation meaning some assessment of the reliability and accuracy of a given method.
2. Issues pertaining to the best way to communicate results to others. This is where additional work is needed if the discipline moves to using the logical approach. There are many options which can be considered and which of these is truly ‘best’ remains to be seen. However, in that regard, the logical approach is similar to, and certainly no worse than, our traditional

approach insofar as we have never fully or properly validated our traditional conclusion wording or scale(s).

On the other hand, as a system of reasoning and for presentation of opinions in the forensic realm, the literature is replete with references that support the use of the logical approach (in one form or another).

A reasonably extensive, but not complete, list of select textbooks and articles is provided at the end of this paper that shows the extensive discussion that has taken place over the years (both in FDE and other areas). It should be noted that, in the course of the discussions represented by those articles, nobody has ever presented a single cogent and sustained counter-argument against this approach.

Advantages/strengths:

The key advantage and strength comes from this approach being founded in logic—literally, it is the application of probabilistic logic. Applying the approach diligently and carefully ensures a logically sustainable result.

Key points:

- It is based on probabilistic reasoning and logic at its core which gives us a clear mechanism to address the omni-present uncertainty in our analyses and reasoning
- It enhances transparency and thoroughness because we must state conditions and assumptions, and clarify potentially ambiguous information
- It focuses on, and answers, the questions of interest to the trier (though they may not realise it)
- It helps to clarify the role of the expert, relative to other parties involved in the judicial decision-making process
- It does not overstep the bounds of science and knowledge

Another advantage is that this approach accommodates quantified data (statistics) very easily. When proper numeric data is available the likelihood-ratio approach becomes feasible, though it is not without challenges. The key issue rests in the proper acquisition of appropriate and relevant data. Unless the data set relates to the propositions in a way that differentiates between them, it will have limited value, or at worst it will completely be useless.

In addition, when dealing the data and statistical estimates the actual mathematics involved can become quite complex due to the conditional nature of the both the dataset and the evaluation (the latter always being conditioned by the framework and the propositions).

On the other hand, such complexity is just a function of the problem, not the solution. It should not be seen as a serious impediment.

Finally, another advantage/strength of this approach is that it works with ALL types of evidence to address any problem of an evaluative nature. That is one of the reasons why it was adopted for all disciplines by ENFSI.

Disadvantages/limitations:

The singular disadvantage for this approach is the lack of understanding by examiners and clients.

A huge, and reasonable, concern for many people relates to limited understanding by the judiciary, legal pundits or lawyers. However, these issues can be addressed in time through research and education for all parties.

Ultimately, there are no practical real limitations or disadvantages to this approach other than unfamiliarity with how it works.

Current ‘users’ of this approach:

Estimating the number of ‘users’ is difficult. At the risk of sounding flippant I will say it is being used, in one form or another, by anyone trying to do our work correctly. But that does not translate into very many people at this time. In fact, the majority of today's audience can rest comfortably in their chairs knowing that the status quo is still safe. At least, for now...

A significant effort to standardize this approach has been made in Europe where ENFSI has published an extensive guide intended for all labs and disciplines. It has been formally adopted in a few select labs but, overall, progress has been slow and very limited. I do not know if the program is on schedule but it was originally intended to take many years.

The pace of change (or non-change) is not surprising to me. After all, the status quo, even when demonstrably flawed, is difficult to change. It is particularly problematic for a body politic, like ENFSI, that must function on the basis of consensus for most things. At any rate, the slow uptake in Europe is, I believe, due to many of the same issues seen on this side of the ocean.

In North America this approach has been discussed in various venues. Elements of it are clearly present in documents produced by the NCFS and even OSAC (though the latter attempt is badly flawed). I should note that there is a draft proposal at the QD committee level however I don't know the status of it following the July 2016 meeting.

In addition, I have heard unofficially that a few laboratories have been exploring this approach but I don't think any have proceeded with implementation yet.

Barriers to implementation:

Inertia and the status quo. I feel that this is primarily due to a commonly-held belief there is no need to change. When coupled with uncertainty about the benefit and value of this approach this becomes a tangible barrier.

At the same time, I firmly believe that FDE's are an intelligent lot. Given the opportunity to learn about it most people quickly understand why this is a better approach. They also ‘get’ the basics of this approach quite easily. However, learning to use it ‘for real’ in casework isn't quite so easy. That takes time and effort.

Indeed, from a pragmatic point-of-view, the biggest barrier is the need for education and information. Training, both theoretical and practical in nature, is needed for practitioners... lots of training. Education is also needed for our clients – particularly the judiciary and lawyers. Of course, such things take time, money and resources.

References supporting the application and use of the logical approach (and variants):

- [Interpreting Evidence: Evaluating Forensic Science in the Courtroom](#) by B. Robertson and G. A. Vignaux (J. Wiley, 1995)

- Statistics and the Evaluation of Evidence for Forensic Scientists, 2nd Ed by C. Aitken and F. Taroni. (John Wiley & Sons Ltd, 2004)
- Expert Evidence and Criminal Justice by M. Redmayne (Oxford Univ Press, 2004)
- Evaluating Scientific Evidence by E. Beecher-Monas (Cambridge Univ Press, 2007)
- Bayesian Networks and Probabilistic Inference in Forensic Science by F. Taroni, C. Aitken, A. Biedermann, and P. Garbolino (John Wiley & Sons, 2006)
- S. Willis, L. Mc Kenna, S. Mc Dermott, G. O' Donnell, A. Barrett, B. Rasmusson, T. Höglund, A. Nordgaard, C. Berger, M. Sjerps, J.J.L. Molina, G. Zadora, C. Aitken, T. Lovelock, L. Lunt, C. Champod, A. Biedermann, T. Hicks, F. Taroni, ENFSI guideline for evaluative reporting in forensic science, available at www.enfsi.eu/news/enfsi-guidelineevaluative-reporting-forensic-science (2015)
- Communicating and Interpreting Statistical Evidence in the Administration of Criminal Justice by C. Aitken, P. Roberts, G. Jackson (Royal Statistical Society's Working Group on Statistics and the Law) – available online at <http://www.rss.org.uk/>
 - No. 1: "Fundamentals of Probability and Statistical Evidence in Criminal Proceedings"
 - No. 3: "The Logic of Forensic Proof: Inferential Reasoning in Criminal Evidence and Forensic Science"
- Data Analysis in Forensic Science: A Bayesian Decision Perspective by F. Taroni, S. Bozza, A. Biedermann, P. Garbolino, and C. Aitken (John Wiley & Sons, 2010)
- Statistics and the Evaluation of Evidence for Forensic Scientists by Aitken, CGG & Taroni, F (John Wiley & Sons, 2004)
- B. Robertson, G.A. Vignaux, Interpreting Evidence: Evaluating Forensic Science in the Courtroom (John Wiley and Sons Ltd, 1995)
- Raymond Marquis, Alex Biedermann, Liv Cadola, Christophe Champod, Line Gueissaz, Geneviève Massonnet, Williams David Mazzella, Franco Taroni, Tacha Hicks, Discussion on how to implement a verbal scale in a forensic laboratory: Benefits, pitfalls and suggestions to avoid misunderstandings. *Science and Justice*, 2016
- Aitken, CGG & Taroni, F (2008) Fundamentals of statistical evidence - a primer for legal professionals. *International Journal of Evidence and Proof*, 12, 181-207
- Aitken, CGG & Lucy, D (2004) Evaluation of trace evidence in the form of multivariate data. *In Jour of the Royal Statistical Society. Series C (Applied Statistics)*, 53. 109–122
- Biedermann, A, Bozza, S & Taroni, F (2008) Decision theoretic properties of forensic identification: Underlying logic and argumentative implications. *FSI*, 177. 120–132
- Biedermann, A, Bozza, S & Taroni, F (2009) Probabilistic evidential assessment of gunshot residue particle evidence (Part I): LR calculation and case pre-assessment using Bayesian networks. *FSI*, 191, 24–35
- Buckleton, JS, Triggs, CM & Champod, C (2006) An Extended Likelihood Ratio Framework for Interpreting Evidence. *Science & Justice*, 46. 69–78
- Champod, C (1995) Edmond Locard – Numerical Standards & "Probable" Identifications. *JFI*, 45. 136–163
- Curran, JM, Triggs, CM & Buckleton, J (1998) Sampling in Forensic Comparison Problems. *IN Science & Justice*, 38. 101–107

- Köller, N, Niessen, K, Riess, M & Sadorf, E (2004) Probability Conclusions in Expert Opinions on Handwriting. Substantiation and Standardization of Probability Statements in Expert Opinions München, Luchterhand.
- Marquis, R, Schmittbuhl, M, Mazzella, WD & Taroni, F (2005) Quantification of the shape of handwritten characters: A step to objective discrimination between writers based on the study of the capital letter O. *FSI*, 150. 23–32
- Marquis, R, Taroni, F, Bozza, S & Schmittbuhl, M (2006) Quantitative characterization of morphological polymorphism of handwritten characters loops. *For Sci Int*, 164. 211–220
- Marquis, R, Taroni, F, Bozza, S & Schmittbuhl, M (2007) Size influence on shape of handwritten characters loops. *For Sci Int*, 172. 10–16
- Stoney, DA. (1991) What made us ever think we could individualize using statistics? *JFSS*, 31(2): 197-199
- Strach, S (1998) Probability Conclusions in Handwriting Comparisons. *IJDE*, 4/4
- Taroni, F, Marquis, R et al. (2012) The use of the likelihood ratio for evaluative and investigative purposes in comparative forensic handwriting examination. *FSI*, 214, 189–194
- Taroni, F, Bozza, S and Aitken, C. (2005) Decision Analysis in Forensic Science. *JFS*, Vol. 50, No. 4
- Cook, R., Evett, IW et al. (1998) A Model for Case Assessment and Interpretation. *Sci and Jus*, 38: 151-156
- Evett, IW. (2000) Verbal Conventions for Handwriting Opinions, *JFS*, v45(2), 508-509
- Strach, SJ (1998) Probability Conclusions in Handwriting Comparisons. *Int'l JFDE*, v4(4), 318-322
- Stockton, A and Day, SP (ASQDE, 2001) Bayes, Handwriting and Science
- Sjerps, MJ, Massier, REF, and Wagenaar, WA (1996) Expressing Expert Opinion Using a Verbal Probability Scale, 5th European Conference for Police and Handwriting Experts, The Hague (Netherlands)
- Strach, SJ (ASQDE 2001) Expressions of Conclusions on Handwriting Comparisons
- Strach, SJ (ASQDE 2002) Conclusions in Forensic Science with Particular Reference to Handwriting Comparisons
- Association of Forensic Science Providers (AFSP), “Standards for the formulation of evaluative forensic science expert opinion”, *Science & Justice*, 2009, 49, 161-164.
- Martire K.A., Kemp R.I., Sayle M., Newell B.R., “On the interpretation of likelihood ratios in forensic science evidence: Presentation formats and the weak evidence effect”, *Forensic Science International* 24 (2014) 61-68.
- I.W. Evett, G. Jackson, J.A. Lambert, S. McCrossan, The impact of the principles of evidence interpretation on the structure and content of statements, *Sci. Justice* 40 (2000) 233–239.
- S.S. Kind, R. Wigmore, P.H. Whitehead, D.S. Loxley, Terminology in forensic science, *J. Forensic Sci. Soc.* 19 (1979) 189–191.
- G.A. Brown, P.L. Cropp, Standardised nomenclature in forensic science, *J. Forensic Sci. Soc.* 27 (1987) 393–399. [8] S.C. Leung, Y.L. Cheung, On opinion, *Forensic Sci. Int.* 42 (1988) 1–13.
- I.W. Evett, Bayesian inference and forensic science: problems and perspectives, *J. R. Stat. Soc. Ser. D (Statistician)* 36 (1987) 99–105.
- G. Jackson, The scientist and the scales of justice, *Sci. Justice* 40 (2000) 81–85.

- C. Champod, I.W. Evett, G. Jackson, J. Birkett, Comments on the scale of conclusions proposed by the ad hoc committee of the ENFSI marks working group, *Inf. Bull. Shoeprint/Toolmark Examiners* 6 (2000) 11–18.
- D. Lindley, The philosophy of statistics, *J. R. Stat. Soc. Ser. D (Statistician)* 49 (2000) 293–337.
- B. Robertson, G.A. Vignaux, C.E.H. Berger, Extending the confusion about Bayes, *Mod. Law Rev.* 74 (2011) 444–455.
- C.E.H. Berger, J. Buckleton, C. Champod, I.W. Evett, G. Jackson, Evidence evaluation: a response to the court of appeal judgment in *R v T*, *Sci. Justice* 51 (2011) 43–49
- D. Faigman, A. Jamieson, C. Noziglia, J. Robertson, R. Wheate, Response to Aitken et al. on *R v T*, *Sci. Justice* 51 (2011) 213–214.
- I.W. Evett, Towards a uniform framework for reporting opinions in forensic science casework, *Sci. Justice* 38 (1998) 198–202.
- C. Champod, I.W. Evett, Commentary on A.P.A. Broeders, 'Some observations on the use of probability scales in forensic identification', *Forensic Linguistics* 6(2): 228–241, *Int. J. Speech Lang. Law* 7 (2000) (1999) 238–243.
- I.W. Evett, Verbal convention for handwriting opinions, *J. Forensic Sci.* 45 (2000) 508–509.
- F. Taroni, C.G.G. Aitken, Fibres evidence, probabilistic evaluation and collaborative test, *Forensic Sci. Int.* 114 (2000) 45–47.
- F. Taroni, J. Buckleton, Likelihood ratio as a relevant and logical approach to assess the value of shoeprint evidence, *Inf. Bull. Shoeprint/Toolmark Examiners* 8 (2002) 15–25.
- F. Taroni, A. Biedermann, Inadequacies of posterior probabilities for the assessment of scientific evidence, *Law Probability Risk* 4 (2005) 89–114.
- A. Biedermann, F. Taroni, C.G.G. Aitken, Letter to the editor — re: conclusion scale for shoeprint and toolmarks examinations, *J. Forensic Identif.* 56 (2006) 685–689.
- W.C. Thompson, E.L. Schumann, Interpretation of statistical evidence in criminal trials: the prosecutor's fallacy and the defence attorney's fallacy, *Law Hum. Behav.* 11 (1987) 167–187.
- I.W. Evett, Avoiding the transposed conditional, *Sci. Justice* 34 (1995) 127–131.
- ENFSI Expert Working Group Marks Conclusion Scale Committee, Conclusion scale for shoeprint and toolmarks examination, *J. Forensic Identif.* 56 (2006) 255–280.
- W.C. Thompson, J. Vuille, A. Biedermann, F. Taroni, The role of prior probability in forensic assessments, *Front. Genet.* 4 (2013) 220 (Article <http://journal.frontiersin.org/article/10.3389/fgene.2013.00220/full>).
- A. Biedermann, F. Taroni, P. Garbolino, Equal prior probabilities: can one do any better? *Forensic Sci. Int.* 172 (2007) 85–93. [34] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A model for case assessment and interpretation, *Sci. Justice* 38 (1998) 151–156.
- G. Jackson, P.J. Jones, in: A. Jamieson, A. Moenssens (Eds.), *Case Assessment and Interpretation*, *Wiley Encyclopedia of Forensic Science*, 2009.
- A. Nordgaard, R. Ansell, W. Drotz, L. Jaeger, Scale of conclusions for the value of evidence, *Law Probability Risk* 11 (2012) 1–24.
- C. Champod, D. Baldwin, F. Taroni, J. Buckleton, Firearm and tool marks identification: the Bayesian approach, *AFTEJ.* 35 (2003) 307–316.
- I.W. Evett, et al., Expressing evaluative opinions: a position statement, *Sci. Justice* 51 (2011) 1–2.

- K.A. Martire, I. Watkins, Perception problems of the verbal scale: a re-analysis and application of a membership function approach, *Sci. Justice* 55 (2015) 264–273.
- M. Sjerps, D.B. Biesheuvel, The interpretation of conventional and Bayesian verbal scales for expressing expert opinion: a small experiment among jurists, *Forensic Linguist.* 6 (1999) 214–226.
- F. Mosteller, C. Youtz, Quantifying probabilistic expressions, *Stat. Sci.* 5 (1990) 2–12.
- K.A. Martire, R.I. Kemp, I. Watkins, M.A. Sayle, B.R. Newell, The expression and interpretation of uncertain forensic science evidence: verbal equivalence, evidence strength, and the weak evidence effect, *Law Hum. Behav.* 37 (2013) 197–207.
- C.Mullen, D.Spence, L. Moxey, A.Jamieson, Perception problems of the verbal scale, *Sci. Justice* 54 (2014) 154–158.
- W. Brun, K.H. Teigen, Verbal probabilities: ambiguous, context-dependent, or both? *Organ. Behav. Hum. Decis. Process.* 41 (1988) 390–404.
- D. McQuiston-Surrett, M.J. Saks, Communicating opinion evidence in the forensic identification sciences: accuracy and impact, *Hast. Law J.* 59 (2008) 1159–1189.
- G.Jackson, I.W. Evett, C. Champod, J. Buckleton, Letter to the editor — re: perception problems of the verbal scale, *Sci. Justice* 54 (2014) 180.
- I.J. Good, Weight of evidence and the Bayesian likelihood ratio, in: C.G.G.Aitken, D.A. Stoney (Eds.), *The Use of Statistics in Forensic Science*, Ellis Horwood, Chichester, UK 1991, pp. 85–106.
- C.G.G. Aitken, F. Taroni, A verbal scale for the interpretation of evidence, *Sci. Justice* 38 (1998) 279–281.
- A. Biedermann, T. Hicks, R. Voisard, F. Taroni, C. Champod, C.G.G. Aitken, I.W. Evett, E-learning initiatives in forensic interpretation: report on experiences from current projects and outlooks, *Forensic Sci. Int.* 230 (2013) 2–7.