A Gap Analysis (Friction Ridge) – Process Improvement

Michele Triplett

Member of the IAI

The 2009 NAS Report *Strengthening Forensic Science in the United States: A Path Forward* [1] highlighted concepts needing improvement for the forensic science disciplines. 16 years have passed with a lot of effort towards improvement. A gap analysis for the Friction Ridge (FR) discipline can assist in identifying areas needing further development which will help new practitioners understand the past, the present, and pending changes for the future. This can also help existing practitioners stay informed about the changes proposed in the discipline. A gap analysis could also give researchers ideas on new areas to examine.

In 2009, it was commonplace for the FR discipline to make unwarranted claims. Since then, the discipline has made vast improvements in the items below:

a) Fingerprints are proven to be unique

- Refuted: Science does not attempt to prove, science provides support
    - Although uniqueness is supported, the amount of information required to determine uniqueness has not been established: a full fingerprint vs. partial prints.

b) Conclusions are absolute and conclusive

- Refuted: The only absolute conclusions in science are when concepts are falsified (e.g. the earth is not flat)
- Scientific conclusions are usually derived by induction or deduction

c) There is a zero-error rate

- No scientific method produces a zero-error rate
- Errors can stem from a deficient method or inappropriate application of a method

There are other concepts that have been improved at a lower level of acceptance:

a) Falsification produces better results than confirmation

b) ACE-V as an acronym, not a method

- Methods have specific rules and parameters
- ACE-V is widely used and accepted; ACE-V is currently being presented as a process in place of a method.

c) *Source Identification* is different than an *Identification*

- The conclusion of *Identification* has been associated with being categorical (a fact), whereas *Source Identification* is the weighing of two propositions

d) *Inconclusive* may be the most warranted conclusion in some situations

    - *Inconclusive* is the warranted conclusion when information is highly ambiguous and/or limited

e) Exclusions based on pattern type, or one discrepancy are not always accurate

    - Science develops tentative conclusions and then tests the tentative conclusions, science does conclude based on one piece of information

This gap analysis is the perspective of the author based on discussions and questions that continue to arise. Gaps are categorized into the following groups: Method, Conclusions, Verification, Error Rates, Documentation and Quality Assurance (including proficiency testing).

Method

- **What is the method?**
    - A method has specific rules and parameters.
    - Both induction and deduction are accepted scientific methods, commonly used within hypothesis testing.
- **What are the parameters within the method?**
    - Are examinations and comparisons physical or analytical tasks? The type of task is important to identify since each has different scientific requirements. Physical tasks require reproducibility (yeast making dough rise should happen regardless of the person performing the task), while analytical tasks require appropriate justification to support the conclusion, to include correct application of the method.
    - What pieces of information are considered Task Relevant (necessary) for a comparison and arriving at a conclusion?
        - Some claim only the questioned and known impressions are necessary, while others feel additional information is necessary (e.g., substrate, matrix, etc.).
    - Does a linear method produce better results than going back and reanalyzing information?
        - Is reanalyzing information the same as circular reasoning?
            - Science promotes reanalyzing information to collect additional data. Circular reasoning is not the reanalyzing of information; circular reasoning is when there is a preconceived conclusion and information is sought that supports the preconceived conclusion, perhaps ignoring data that does not fit the preconceived conclusion.
    - What information within impressions can be used?
        - Can incipient ridges, scars, pores, warts, ridge edges, etc. be used as support for an identification? Currently, each forensic service provider determines the information that can be used within their agency to support a source conclusion as noted by the Academy Standards Board, Standards Development Organization, "The FSP shall define the observed data examiners use during the examination of friction ridge impressions" [2].
        - In addition to feature type, ridge counts, spatial relationship between features, clarity and quantity are factors to consider during a comparison.
    - How is it decided that a difference between images is categorized as either a dissimilarity or disagreement?
        - What are the distortional clues to determine artifacts over real features, for both the questioned and known impressions (such as livescan anomalies)?
    - Is there a PPI requirement for digital images?
        - While some claim a requirement is necessary, others claim that PPI is regarding the resolution, not the quality of an image.
    - The amount of information required is discussed in the conclusions section below.
- **Is deduction an acceptable method within the FR discipline?**
    - A deductive conclusion is when an impression is compared and identified to person A and therefore, it can be deduced that all others can be excluded (a theoretical exclusion). Of course, the strength of this type of conclusion is only as strong as the initial identification arrived at through the induction method.
    - Many examiners have stated that source conclusions can only be derived through a direct comparison of impressions (i.e., induction), source conclusions arrived at through deduction is not a topic that has not been thoroughly discussed or debated yet.

- Weighing propositions is a method for determining the amount of support for a conclusion. The amount of support then leads to a conclusion under the Bayesian approach for arriving at conclusions. Currently, the only means to weigh support is by considering examiners' confidence levels.

- **Is confidence an acceptable means of weighing information?**
  - Relying on confidence is a highly subjective measurement. Should the FR discipline attempt to increase objectivity?
- **Is difficulty the same as complexity?**
  - Difficulty is usually a subjective assessment and may be different for each person based on their abilities, while complexity should be considered an objective measure of technical difficulty of tasks. The two concepts need to be differentiated, regardless of the terminology used.
- **Should examiners speculate on the activity performed during the touching of an object due to the orientation of a questioned impression (e.g., proposing a bottle was handled in a way that could have been used as a weapon)?**

Conclusions

- **What are the possible conclusions?**
  - It is standard knowledge that agencies have a differing number of possible conclusions; some use three conclusions, some use five conclusions, and some may use nine or ten conclusions. Is one option better than the others?
- **How much information is needed to arrive at a conclusion?**
  - It is common knowledge that a feature count does not accurately represent the strength of a conclusion because the feature type, clarity and quantity of information is just as important as the number of features.
- **Is there a difference between opining, deciding and concluding?**
  - Do practitioners know the difference between the meanings of these words?
  - Should language be used that is understandable to the courts?
- **IDENTIFICATION (or SOURCE IDENTIFICATION)**
  - Is Source Identification different than Identification?
    - The conclusion of *Identification* has been associated with being categorical (a fact), whereas *Source Identification* is the weighing of two propositions.
  - Should Identifications be measured for strength (e.g. the quality and quantity of a rolled tenprint vs. the Mayfield impression [3])?
  - Should Identifications stand on their own or does a comparison with less consistency warrant an Identification if a person has already been Identified several times on an object or in a case?
- **AFIS / ABIS IDENTIFICATIONS**
  - Is a higher threshold needed for AFIS candidates due to the potential for close non-matches?
    - Tenprint impressions are searched and identified worldwide each day. The error rate for AFIS tenprint identifications are a fraction of AFIS latent identifications, demonstrating that all AFIS identifications are not prone to error.
    - The Alton Dandridge error [4] did not stem from an AFIS search. Does this show that all Identifications, including manual comparisons, need a strong threshold?
  - Should thresholds be higher for all complex Identifications, from both manual and AFIS comparison, once complexity can be measured?
- **INCONCLUSIVES**
  - Is there an issue with some examiners concluding inconclusive when an identification is warranted? How is it determined that an identification is warranted if there are no parameters for an identification?
- **EXCLUSIONS**
  - Are exclusions categorical (facts) or inferences?
    - In science, the only conclusive conclusions are when something is falsified.

- Are erroneous exclusions an acceptable limitation of the comparative sciences?
- Once a questioned impression has been identified, can all other known exemplars be theoretically excluded by the use of logical deduction?

Verification

- **Should verification be an independent assessment to determine reproducibility, or should verification be a review to ensure the application of the method was correct?**
- **Can AFIS make Identifications and have a human as a verifier?**
  - Consider lights-out, which has become an acceptable means of arriving at an identification. Is verification necessary for lights-out comparisons?
- **Is Blind verification a higher level of quality control?**
  - Blind verification can only determine reproducibility of the conclusion, not determine if the application of the method is correct.
- **Should multiple forms of verification be utilized? What combination of practices are the best and why?**

Error Rates (ERs)

- **Are errors inevitable or can parameters be implemented to prevent errors?**
- **Are ER studies testing the method or the practitioner?**
  - ER studies have not tested a specific method; they have tested examiners ability to arrive at ground truth conclusions.
  - Study results may/may not be representative of a specific method for arriving at conclusions.
- **Are erroneous conclusions the only type of error?**
  - In science, it is an error to arrive at a correct conclusion inappropriately (e.g., performing long division incorrectly but arriving at the correct conclusion by chance).
  - Should ERs consider the application of a method and not simply the conclusion?
- **Are current ER studies representative of errors in casework?**
  - ER studies compare examiner conclusions to ground truth conclusions, a detail that is not known in casework. ERs from studies may/may not be extrapolated to give an indication of errors in casework.
  - The possible conclusions in studies may not be the same as the options examiners use in casework.
  - The Hawthorne Effect may make ERs in studies different than ERs in casework [5].
  - Is the complexity of the comparisons in research studies representative of the complexity of comparisons in casework? This is unknown since there is not a standardized means of measuring the complexity of a comparison. In studies, examiners have been asked about the difficulty, not necessarily about the complexity. If the complexity level has been asked, the rating has been a subjective rating.
- **Should ERs consider the risk of error in a specific conclusion?**
  - Current ERs look at overall errors within the discipline which may artificially portray the strength of a conclusion that has weaker support.
  - An overall ER may/may not be representative of a specific comparison and conclusion.
- **What measures can be implemented to identify and reduce errors?**
  - Types of errors need to be identified before implementing methods for reduction, (e.g., is inappropriate support for a conclusion an error?).
  - Current recommendations to minimize bias and improve conclusions include masking certain information and/or documentation of confidence levels.

- Close non-matches are discussed within the Quality Assurance section.
    - If a conclusion is close, how is it known that it is a non-match?

Documentation

- **Are checklists that state attributes of a questioned impression (matrix, substrate, clarity, etc.) valuable in helping another person determine what was done in a case?**
    - Are the attributes listed on a checklist measurable, are the ratings assumptions or generalizations based on the opinion/experience of the analyst?
- **Should the analysis of the known exemplars be documented and include checking the correct sequence and listing any false features due to distortion prior to a comparison being performed?**
- **Should documentation of the support for the conclusion be the necessary type of documentation desired?**

Quality Assurance

- **Are Quality Systems in place to identify and resolve issues?**
    - Agency problems and errors have been identified by sources outside of agencies.
    - Problems found externally indicate an internal quality system lacks the ability to identify areas needing improvement. The root cause and corrective action should be regarding the quality system, which is rarely identified as the root cause of problems found.
        - For example, in the FBI Mayfield error, "the OIG identifies as the primary factor for the mistake made by the FBI and the defense experts the extraordinarily "unusual similarity" between the two prints" [6]. Many practitioners find this cause inappropriate, believing the cause was due to overinterpretation of information. Is 'unusual similarity' or 'close non-match' being used as justification for errors? Misleading causes do not assist in finding corrective actions that will improve the FR discipline.
- **How can procedures be streamlined to eliminate wasted efforts instead of doing more with less?**
- **Is the discipline transparent about deficiencies and improvements that are needed?**
    - Is it transparent to label a discipline as a forensic science, yet not adhere to scientific requirements?
    - Is it transparent to use influential words that may overstate concepts (such as the use of the word *conclusion* for an *opinion*)?
        - A conclusion requires a basis for the result, whereas an opinion does not.
    - Are Best Practices really the best if they are not compared to other practices?
    - Is it transparent to imply probabilities are used when the probabilities are personally derived?
- **Are proficiency tests adequate?**
    - Is the complexity of comparisons on proficiency tests adequate?
    - Should proficiency tests include the ability to determine if an impression is of value for a comparison?
    - Should proficiency tests include the ability to correctly determine orientation of a questioned impression?
    - Should proficiency tests include the ability to recognize tonally reversed impressions?
    - Should proficiency tests include the ability to determine distortional artifacts over real features, and the observational clues for the distortion?
    - Should proficiency tests include the ability to apply the method correctly?
    - Should proficiency tests include the ability to arrive at inconclusive conclusions when inconclusive is the best supported conclusion?
    - Should proficiency tests include the ability to verify a case and identify deficiencies?

- Does a score below 100% indicate an examiner is not proficient?

- **Are proficiency tests consistent in how they are administered?**
  - Are the conclusions used on proficiency tests the same conclusions that are used in casework?
  - If tests are administered differently, comparisons of test results can easily be misconstrued.
- **Is the discipline able to determine if bias is a factor, or is it being assumed?**
  - Bias is more prevalent when information is vague and/or limited. The lack of specificity for a method promotes the ability for bias to be possible.
  - Can bias be positive in that it informs you of information you were not aware of, or is bias always negative because it is a reliance on irrelevant information?
- **Are standards intentionally vague to allow for flexibility?**
  - Flexibility is valuable; however, too much flexibility creates chaos, inconsistency and a lack of standardization.
- **When is consensus considered *false consensus* [7]?**
  - Is consensus equivalent to validity?
- **Do we value discussion, debate and differing views?**
  - Are questions discussed, or are alternative views disregarded?
- **Do early adopters recognize the value of suggestions faster than others, or are they deferring to authority?**
- **How should an agency develop an appropriate staffing model?**
  - Many agencies determine the number of staff hired by their budget, instead of evaluating the workload and expected production level from each staff member. The production level is different for each agency based on the tasks required.

There are an abundance of issues within the Friction Ridge discipline to research and clarify. This list can and should be expanded as issues are noted. Many of these issues are dependent on having a defined method. If the desire is to keep the method as flexible as possible, then consistency in conclusions is not likely. Minimally, the use of different methods should be compiled so different methods can be compared and assessed for their benefits and limitations. This gap analysis will give researchers and those creating standards areas to investigate that may not have been previously considered. Addressing any one of these issues would be a great benefit to the discipline.

References

[1] National Research Council. *Strengthening forensic science in the United States: A path forward*. Washington, D.C., The National Academies Press, 2009.

[2] ANSI/ASB Friction Ridge Consensus Body. *Standard 015, Standard for Examining Friction Ridge Impressions,* 1st Ed., 2024.

[3] Office of the Inspector General. *A Review of the FBI's Handling of the Brandon Mayfield Case.* https://oig.justice.gov/sites/default/files/archive/special/s0601/Chapter4.pdf, 2006

[4] Double Loop Podcast. *Beniah Dandridge Case – Fingerprint Comparisons.*

https://doublelooppodcast.com/2018/10/beniah-dandridge-case-fingerprint-comparisons/. Accessed 5/18/25.

[5] Landsberger, H. A. *Hawthorne Revisited: Management and the Worker, Its Critics, and Developments in Human Relations in Industry.* N.Y.S. School of Industrial and Labor Relations, Cornell University, Ithaca, New York, 1958.

[6] Office of Inspector General. *A Review of the FBI's Handling of the Brandon Mayfield Case.*

https://oig.justice.gov/sites/default/files/legacy/special/s0601/final.pdf, page 322. March 2006.

[7] Ross, Lee; Greene, David; House, Pamela. *"The 'false consensus effect': An egocentric bias in social perception and attribution processes". Journal of Experimental Social Psychology.* ***13*** *(3): 279–301*. doi:10.1016/0022-1031(77)90049-X, May 1977.