B165 Black Box and White Box Forensic Examiner Evaluations — Understanding the Details

R. Austin Hicklin, MS, 3150 Fairview Park, Falls Church, VA 22042; Bradford Ulery, BA, 3150 Fairview Park Drive, S, Falls Church, VA 22042-4519; Maria A. Roberts*, 2501 Investigation Parkway, Quantico, VA 22135; and JoAnn Buscaglia, PhD*, FBI Laboratory, CFSRU, 2501 Investigation Parkway, Quantico, VA 22135*

The goal of this presentation is to assist attendees in understanding the issues involved in using the black box approach to assess the accuracy and reliability of forensic examiners.

This presentation will impact the forensic science community by describing an approach for assessing potential areas of strength and weakness in forensic science, as well as by offering several objective measures to support admissibility requirements.

Recently, there has been a great deal of interest in using "black box and "white box" techniques to evaluate decisions made in a variety of forensic disciplines. This presentation will help dissect the details of conducting such evaluations, which are not as deceptively simple as they seem.

Black box evaluations are conducted by assessing the examiner's decision without regard to how those decisions are made. Black box evaluations can provide a useful overall understanding of the accuracy, reproducibility, and repeatability of the decisions made in response to a given task. Such evaluations do not attempt to assess how a specific examiner performs on specific data — but black box evaluations are a necessary first step toward such detailed tests. Black box evaluations provide a means of quantifying forensic examinations for which quantitative models do not (yet) exist and, therefore, provide both an interim solution while such models are under development, as well as a means of validating such models.

Conversely, white box evaluations are conducted to gain an understanding of how and why examiners make decisions. White box evaluations are detailed assessments of the bases of examiners' decisions, focused not just on the end decisions but on the features and attributes used by the examiners in rendering conclusions. While analyses of black box results deal with the inter-examiner variability of decisions, white box analyses also deal with inter-examiner variability of the detection of features and other attributes.

This presentation will discuss topics that should be considered in the design of black and white box evaluations, including: (1) representativeness of data (dealing with heteroscedastic data, avoiding biased data selection); (2) assessing accuracy vs. reproducibility and repeatability (methods of measurement, data selection implications); (3) test size (precision of measurement, measuring rare events); (4) the Hawthorne effect (dealing with the differences between behavior in tests vs. operations, minimizing differences between test and operational procedures); (5) measuring rates of errors and non-consensus decisions; and, (6) dimensions of examiner skill (accuracy and effectiveness).

Evaluation, Error Rates, Examiner Accuracy