

Measuring the accuracy of facial forensics comparisons

FOUR STANDARD SURVEY QUESTIONS

1. Explain who will be surveyed and why the group is appropriate to survey.

In the paper *Perceptual expertise in forensic facial image comparison* by White, Phillips, Hahn, Hill, and O’Toole, the perceptual accuracy of facial forensic experts was measured, and examiners were shown to be more accurate than the general public. However, the study did not explain why. A follow-up study will be conducted to measure the accuracy of facial forensic examiners. The study will have two parts. The first part of the collection will measure the accuracy of forensic examiners using their tools and laboratory methods. The second part of the collection will survey participants on their background, experience, and education. In addition to facial examiners, there will be two control groups: fingerprint examiners and non-examiner face experts. Participants for the two control groups will be selected because they have backgrounds and experiences similar to facial examiners.

There are three collection instruments for this survey. The first instrument is a screening questionnaire, the second instrument is a background questionnaire, and the third is an image-pair comparison questionnaire. There are three background questionnaires, but each group will only complete one background questionnaire. Additional details about the instruments are given in the answers to Questions 3 and 4. All participants will take the screening questionnaire, responses will be recorded by the interviewer and will be used to determine if a volunteer is eligible to participate in the study. Responses received from the background questionnaire will be used to determine which factors contribute to greater accuracy in facial comparison. Responses for the background questionnaire will be saved by the web-based survey.

All eligible participants in the study will be asked to compare 20-30 pairs of face-images. After comparing each pair of face-images, participants will be asked to answer two questions that are contained in the comparison questionnaire. The responses to the comparison questions will be used to measure the accuracy of the participants in performing facial comparisons.

The three groups in this study will be facial forensic examiners, fingerprint examiners, and non-examiner face experts. All three groups will be surveyed by answering two questionnaires.

2. Explain how the survey was developed including consultation with interested parties, pre-testing, and responses to suggestions for improvement.

The background questionnaires to be administered are based on previous studies as well as consulting with experts in facial forensics and experts in survey design at NIST. These questionnaires are based on the survey used to assess the backgrounds of forensic fingerprint examiners in the paper *Accuracy and reliability of forensic latent fingerprint decisions* by Ulery,

Hicklin, Buscaglia, and Roberts. The background questionnaires are designed to help answer questions that arose from the results in the paper *Perceptual expertise in forensic facial image comparison* by White, Phillips, Hahn, Hill, and O'Toole.

The comparison questions are based on the paper by White, Phillips, Hahn, Hill, and O'Toole, and consulting with experts in facial forensics.

The screening questionnaire was based on consultations with experts in forensics and the previous study *Accuracy and reliability of forensic latent fingerprint decisions* by Ulery, Hicklin, Buscaglia, and Roberts.

3. Explain how the survey will be conducted, how customers will be sampled if fewer than all customers will be surveyed, expected response rate, and actions your agency plans to take to improve the response rate.

Participants will come from the facial forensic, face recognition, and fingerprint communities. Emails asking for potential participants will go to professional groups such as the Facial Identification Scientific Working Group (FISW1G), the Facial Identification subcommittee of the Organization of Scientific Area Committees (OSAC), the European Network of Forensic Science Institutes (ENSFI), and the Scientific Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST).

The questionnaires are part of a study to measure the accuracy of three groups that will be performing facial forensic comparisons. After reviewing the study and consent form, volunteers who choose to participate will sign the consent form and proceed to the screening questionnaire. Participants will be asked the 12 screening questions over the phone; four of these are concerning contact information in order to send materials (such as the background questionnaire and images) and to be sure appropriate precautions are in place when sending or receiving information via fax. As stated in the screening questionnaire, a postal address is needed to send a participant's accuracy score to them if requested. However, the participant may choose to defer providing this information until they decide if they wish to request their score. The screening questionnaire is used only for contact information and to determine if the volunteer fits into one of the three groups.

After it is determined if the participant fits into one of the three groups, the participant will be given a link to the group-appropriate background questionnaire, which will be taken online where the answers will be saved by the web-based survey. An individualized link to the background questionnaire will be emailed to the participant. There are three background questionnaires, but each questionnaire is only for a specific group. Every participant will complete the same background survey as every other participant in their group, but the individualized URL will allow linking of responses to participant. Participants will be given a link to a web-based survey for answering the comparison questions. For the comparison questions, an individualized URL will allow linking of responses to participants.

We estimate that we will receive 75 responses from each of the three groups described above.

Facial forensics examiners are people who perform forensic facial comparisons as part of their jobs and have been trained to perform facial forensic comparisons.

Fingerprint examiners are people who perform forensic fingerprint comparisons as part of their jobs and have been trained to perform fingerprint comparisons. These examiners must not perform or have performed facial forensic comparisons as part of their jobs. They must also have been trained to perform facial forensic comparisons.

Non-examiner face experts are people who are knowledgeable about face recognition. These people neither are nor have been forensic examiners. They also have not been trained in how to perform forensic comparisons. However, they have a combination of experience conducting research in face recognition, managing face recognition programs, or addressing policy issues related to face recognition.

All respondents participating are expected to complete the screening questionnaire as they have agreed to take part in the study. We expect all participating respondents to complete the background questionnaire.

PII will be collected for administering the study. The respondents' personally identifiable information (PII) will be protected in the following manner. Each subject enrolled in this study will be assigned a study number and this number will identify all data, including results of facial comparisons. A master list linking the PII to the study number will be kept on an encrypted disk located in a locked office or a locked cabinet at NIST. Responses to the screening and background questionnaires will be treated as PII and will not be shared outside of NIST. Personal privacy information would be withheld from release to the public under the Freedom of Information Act, 5 U.S.C. 552, exemption (b)(6): personnel and medical files and similar files the disclosure of which would constitute a clearly unwarranted invasion of personal privacy.

4. Describe how the results of the survey will be analyzed and used to generalize the results to the entire customer population.

This study will address concerns raised in the National Academy of Sciences (NAS) forensic study *Strengthening Forensic Science in the United States: A Path Forward*. The NAS study called for more scientific rigor and the need to establish an underlying forensics science, particularly in forensic pattern analysis such as comparing two faces. Facial forensic examiners are trained to compare faces that appear in images and videos to determine if the two faces are the same person or not, providing justification for their conclusions. They are prepared to testify in court and to defend their conclusions.

The first aim of the study is to measure the accuracy of the three groups at comparing faces. The answers from the comparison questions will allow us to estimate the accuracy of the three groups.

Another aim of this study is to understand which factors in an examiner's background contribute to greater accuracy. Answers to the background questionnaires will be used to characterize the backgrounds of facial forensic examiners, fingerprint examiners, and non-examiner face experts.

Analyzing the responses from all three groups will assist in identifying factors that contribute to the accuracy of facial examiners.

The resulting understanding of the sources of facial examiner accuracy will help in establishing a forensic science. In turn, this will help in meeting the Daubert standard for facial examinations in U.S. Courts. The Daubert standard is used to determine if expert evidence is admissible.

Using methods from signal detection theory, analysis of variance, and correlation theory, the backgrounds of these three groups will be compared with results from the accompanying experiment measuring face recognition accuracy. This analysis will be the first to study the background of facial examiners and their accuracy.

National Research Council. (2009). *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: The National Academies Press.

Ulery, B., Hicklin, R., Buscaglia, J., & Roberts, M. (2011). Accuracy and reliability of forensic latent fingerprint decisions. *Proceedings of the National Academy of Sciences of the United States of America*, 108(19), 7733-7738. doi:10.1073/pnas.1018707108

White, D. Phillips, P.J., Hahn, C.A., Hill, M., & O'Toole A.J. (2015). Perceptual expertise in forensic facial image comparison. *Proceedings of the Royal Society of London B*, 282(1814). doi:10.1098/rspb.2015.1292