OSAC 2022-S-0001 Standard Guide for Image Comparison Conclusions/Opinions

Digital/Multimedia Scientific Area Committee
Organization of Scientific Area Committees (OSAC) for Forensic Science



OSAC Proposed Standard

OSAC 2022-S-0001 Standard Guide for Image Comparison Conclusions/Opinions

Prepared by
Digital/Multimedia Scientific Area Committee
Version: 2.0
June 2023

Disclaimer:

This OSAC Proposed Standard was written by the Organization of Scientific Area Committees (OSAC) for Forensic Science following a process that includes an open comment period. This Proposed Standard will be submitted to a standards developing organization and is subject to change.

There may be references in an OSAC Proposed Standard to other publications under development by OSAC. The information in the Proposed Standard, and underlying concepts and methodologies, may be used by the forensic-science community before the completion of such companion publications.

Any identification of commercial equipment, instruments, or materials in the Proposed Standard is not a recommendation or endorsement by the U.S. Government and does not imply that the equipment, instruments, or materials are necessarily the best available for the purpose.

To be placed on the OSAC Registry, certain types of standards first must be reviewed by a Scientific and Technical Review Panel (STRP). The STRP process is vital to OSAC's mission of generating and recognizing scientifically sound standards for producing and interpreting forensic science results. The STRP shall provide critical and knowledgeable reviews of draft standards or of proposed revisions of standards previously published by standards developing organizations (SDOs) to ensure that the published methods that practitioners employ are scientifically valid, and the resulting claims are trustworthy.

The STRP panel will consist of an independent and diverse panel, including subject matter experts, human factors scientists, quality assurance personnel, and legal experts, which will be tasked with evaluating the proposed standard based on a comprehensive list of science-based criteria.

For more information about this important process, please visit our website at: https://www.nist.gov/topics/organization-scientific-area-committees-forensic-science/scientific-technical-review-panels.

Standard Guide for Image Comparison Conclusions/Opinions[1]

1. Scope

1.1 This standard defines conclusion (hereafter "opinion") categories that shall be reached by a practitioner performing comparisons of people, objects, or scenes captured in images (e.g., faces, vehicles, clothing, skin details), regardless of the process by which opinions are reached (i.e., the examination methodology).

1.1.1 This standard does not assign quantitative probability values, including numerical assessments of the strength or weight of evidence, to opinion categories. If quantitative probability values are used by an examiner, they shall be reported with these categories.

1.1.2 This standard does not address how to form, document, or report an opinion (FISWG Minimum Guidelines for Facial Image Comparison Documentation, SWGDE Technical Overview for Forensic Image Comparison).

1.1.3 This standard does not supersede published opinion scale standards for other disciplines.

1.1.4 This standard is based upon practical experience, research, and resources available at the time of publishing. Published research[2][3][4] demonstrates that trained practitioners

[1] There is a movement in the forensic community to eliminate the word "conclusion" from the formal set of words that describe forensic processes. For example, ISO does not use the word "conclusion". This is reflected by the Organization of Scientific Area Committees for Forensic Science (OSAC) preference to use the term "opinion" (defined as View, judgment, belief – takes into consideration other information in addition to observations, data, calculations, and interpretations).

[2] Phillips, P. , White, D. , O'Toole, A. , Hahn, C. and Hill, M. (2015), Perceptual expertise in forensic facial image comparison, Proceedings of the Royal Society B-Biological Sciences, [online], https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=917560 (Accessed March 21, 2022)

[3] Phillips, P. Jonathon & Yates, Amy & Hu, Ying & Hahn, Carina & Noyes, Eilidh & Jackson, Kelsey & Cavazos, Jacqueline & Jeckeln, Géraldine & Ranjan, Rajeev & Sankar, Swami & Chen, Jun-Cheng & Castillo, Carlos & Chellappa, Rama & White, David & O'Toole, Alice. (2018). Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. Proceedings of the National Academy of Sciences. 115. 201721355. 10.1073/pnas.1721355115.

[4] Bruehs WE, Tucker NM, Meline KA. Observer determination of the make, model, and year of questioned vehicles. J Forensic Sci. 2022;67:200-206. https://doi.org/10.1111/1556-4029.14944

are effective in image comparison. At present, there is limited research[5] that addresses the ability of practitioners to reproducibly apply the opinion categories listed in this standard or the ability of laypersons to interpret their meaning.

1.2 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

1.3 *This standard is intended to be used by individuals with discipline specific knowledge, skills, and abilities acquired through education, training, and experience.*

2. Referenced Documents

2.1 *ASTM Standards*:

2.1.1 Standard Guide for Facial Image Comparison Feature List for Morphological Analysis (E3149)

2.2 *FISWG Standards:*

2.2.1 FISWG Minimum Guidelines for Facial Image Comparison Documentation

2.3 *SWGDE Standards:*

2.3.1 SWGDE Best Practices for Photographic Comparison for All Disciplines

2.3.2 SWGDE Technical Overview for Forensic Image Comparison

3. Terminology

3.1 *Definitions specific to this standard:*

[5] Hahn, C. , Tang, L. , Yates, A. and Phillips, P. (2021), Forensic facial examiners vs. super-recognizers: Evaluating behavior beyond accuracy, PsyArXiv Preprints; OSF, [online], https://doi.org/10.31234/osf.io/hq2ab, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=932786, https://psyarxiv.com; osf.io (Accessed March 21, 2022)

3.1.1 *source*, *n*—the subject matter captured in an image

3.1.1.1 *Discussion – Examples of a subject matter include a person, object, or scene.*

4. Significance and Use

4.1 This standard is intended to increase harmonization and consistency by providing a framework of opinion categories across and within forensic image comparison disciplines. At the time of publication, standardized and validated opinion scales do not exist for many disciplines performing source determinations of people, objects, or scenes depicted in images.

4.2 . Each discipline should create and validate opinion scales based on the source opinion categories in this standard.

4.2.1 In the absence of discipline specific opinion scales, operational opinion scales should be built using these categories.

4.3 Validation shall include reference to any empirical studies of the opinion scale applied to a given type of evidence and interpretation.

4.3.1 In the absence of studies, the opinion scale should explicitly state that no such studies exist.

4.4 To ensure interorganizational consistency, this framework categorizes opinions based on the level of support provided by the data for the propositions under consideration. When defining opinion scales, organizations shall consider alternative propositions (both the propositions of same source and of different source).

4.5 The labels of the opinion scale may differ from those listed in this standard, but they should explicitly correspond to the opinion categories defined in this standard.

4.5.1 The opinion categories "Support for Exclusion" and "Support for Common Source" may be subdivided into more specific levels of relative support when empirical research demonstrates that examiners can accurately and reliably gauge the more finely grained categories.

4.6 Image comparison is a subjective practice in nature. Organizations should ensure appropriate procedures are in place to promote consistent implementation of their opinion scales.

5. Opinion Categories

5.1 ***Exclusion:*** an opinion that the observed dissimilar characteristics far outweigh the observed similar characteristics or where no distinctive similarities are observed. The observed characteristics are much more probable given the proposition that the images depict a different source, rather than the proposition that they depict the same source.

5.2 ***Support for Exclusion***: an opinion that the observed dissimilar characteristics outweigh the similar characteristics but are insufficient to reach an exclusion. The observed characteristics are more probable given the proposition that the images depict a different source, rather than the proposition that they depict the same source.

5.3 ***Inconclusive***: an opinion that the observed characteristics are equally probable given the proposition that the images depict different sources and the proposition that the images depict the same source.

5.4 ***Support for Common Source***: an opinion that the observed similar characteristics outweigh the observed dissimilar characteristics but are insufficient to reach strong support for common source. The observed characteristics are more probable given the proposition that the images depict the same source, rather than the proposition that they depict a different source.

5.5 ***Strong Support for Common Source***: an opinion that the observed similar characteristics far outweigh the observed dissimilar characteristics. The observed characteristics are much more probable given the proposition that the images depict the same source, rather than the proposition that they depict a different source.

5.5.1 A practitioner shall confine an opinion to the support-based categories provided in this guide and shall not opine that two items (e.g., faces, vehicles, clothing, skin details) definitively (e.g., to the exclusion of all others) originate from the same source . A practitioner shall not use terms in the stated opinion such as ‘individualize’ or ‘individualization.’ A practitioner shall not express an opinion as an absolute certainty.

APPENDIX[6] (NON-MANDATORY CONTENT)

X1. OPINION SCALE EXAMPLES

X1.1 Example Opinion Scale 1

X1.1.1 Exclusion = Exclusion

X1.1.2 Inconclusive = Inconclusive

X1.1.3 Investigative Lead = Support for Common Source

X1.2 Example Opinion Scale 2

X1.2.1 Non-Ident = Exclusion

X1.2.2 Unlikely = Support for Exclusion

X1.2.3 Inconclusive = Inconclusive

X1.2.4 Likely = Support for Common Source

X1.2.5 Highly Likely = Strong Support for Common Source

[6]Examples contributed from OSAC Facial Identification and Video/Imaging Technology & Analysis Subcommittees