DM SAC Task Group Response to STRP Report 2022-S-0001 Standard Guide for Image Comparison Conclusions/Opinions

In the STRP Report 2022-S-0001 Standard Guide for Image Comparison Conclusions/Opinions, there were two views expressed. While view 1 expressed some concerns of the document and indicated some room for improvement, it generally supported the document overall. Alternatively, view 2 did not support the document and provided multiple areas of concern all of which are listed under section 1 (Scientific and Technical Merit). The DM SAC Task Group appreciated the process of the STRP to include the final report and the dissenting views. There are some items from the report that the DM SAC Task Group made edits to the document and some items that they did not. This reply/response to the STRP final report is meant to provide additional insight on why we did or did not make edits as recommended or requested from the STRP.

View 1:

Some of the specific concerns/room for improvement from view 1 are under section 2 (Human Factors), 3 (Quality Assurance), and 4 (Scope and Purpose). It states that there should be specific statements on quality assurance and human factors not being covered in the document in the scope section. The DM SAC Task Group feels that adding all areas that there are many other areas that are not covered in the document and to list all "out of scope" areas would make the scope larger than necessary. However, we did make a change in the scope section (1.1.2) to state "This standard does not address how to form, document, or report an opinion..." to better clarify that there are many other parts of the process that are not covered in this document.

Additionally, view 1 states under section 5 (Terminology) of the STRP report that "Similarly, some STRP members recommend that "came from the same source" should be explicitly prohibited as a separate statement from "came from the same source to the exclusion of all others." The DM SAC Task Group has made a change to the sentence to now read "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."

View 2:

Under section 1 (Scientific and Technical Merit), view 2 expressed a different critique that the DM SAC Task Group would like to specifically provide a response to certain portions of it.

The first main area of concern that view 2 brought up was on the "validity of the scale", specifically that the lack of cited research on validating the scale. On page 5 of the report, view 2 makes three statements regarding the "scale" stated in the document.

1. " The document implicitly acknowledges the absence of data on repeatability and reproducibility and the absence of any research on implementing the 5-point scale when it states that "[o]rganizations should ensure appropriate procedures are in place to promote consistent implementation of their opinion scales" but provides no guidance on how this should be done."
2. "The document also acknowledges the importance of validating scales when it states that "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." We agree further refinement would require empirical research, just as we respectfully suggest that prior to being placed on the OSAC Registry the proposed scale (or any other scale) needs the same."

3. "If the FSSB determines this standard should be placed on the Registry, the standard should include an unequivocal statement that the scale has not been validated (and that no scale has in this field)."

The DM SAC Task group agrees with the view on validated scales. However, the document is not specifying one scale but is providing "opinion categories" that disciplines should use to write their discipline specific opinion scale document. Because the document discusses categories, validating opinion scales and providing guidance on how to validate a scale is outside the scope of this document. However, the DM SAC Task Group appreciates and supports the great concern that view 2 raises regarding validation for each individual discipline opinions scales. The building blocks of creating and validating the discipline-specific opinion scale lies in individual organizations creating and validating (through intraorganizational and interorganizational research) opinion scales from the opinion categories listed in this document.

The DM SAC Task Group has made some edits within section 4 (Significant and Use) to better clarify the purpose of the document as a framework and to clarify that discipline specific opinion scales should be created and should specifically reference validation studies or the lack thereof.

The second main concern of view 2 was in the presentation of hypotheses, specifically regarding the confusion between a hypothesis and a decision. The DM SAC Task Group agrees. To remedy this, the word "exclusion" was removed from any context in which a hypothesis was intended (section 4.4). "Exclusion" and "Support for Exclusion" are indeed intended to be opinion decisions.

The third main concern of view 2 is in regard to the differences between using a two-stage approach, a Bayesian posterior probability, or a simple likelihood ratio. The document does not attempt to mandate any particular interpretation method as there would unlikely be a consensus at this time. Conforming to the proposed guidelines in this document will generate the practical experience that will be a foundation for future consensus.