



To: Independent Review Panel
From: Legal Resource Committee
Subject: Appeal from Adjudication of LRC Comments on ASTM E3085-17
Date: June 28, 2018

Introduction

The OSAC Legal Resource Committee (LRC) unanimously concluded that ASTM E3085-17 is not suitable for inclusion on the OSAC Registry of Approved Standards.¹ The subject of this appeal is the response—or rather, the lack of a response—to the section of the LRC comments asserting that “Guidelines on Ascertaining ‘Meaningful Differences’ Are Missing.”² The essence of the subcommittee’s reply is that “[t]he guidance is explicit that acquired knowledge of the material and the technique must coexist to best evaluate the data.”

We are perplexed by this response. Because ASTM E3085-17 contains virtually no criteria for ascertaining when a pair of spectra can be distinguished from one another, there has been no adjudication of the committee’s concern³ that the standard does not specify the conditions that justify a determination that a pair of spectra are so different as to “indicate[] the two samples do not share a common origin.” (ASTM E3085-17 § 3.1.6.) The subcommittee has yet to adjudicate the LRC’s comment so as to offer the legal community an explanation of why a standard that directs analysts to decide whether “meaningful differences” are present omits explicit criteria to objectively ascertain which differences are meaningful. The issue of standardless, highly subjective comparisons of feature sets is receiving increased attention by judges and lawyers.⁴ It is vital for OSAC to address this issue in the process before placing standards on the registry. Because we do not believe the issue has been addressed for this standard, we respectfully file this appeal.

Basis for the Appeal

ASTM E3085-17 repeatedly asks analysts to decide if there are “meaningful differences” without giving them any objective standard for doing so. For example:

¹ Kent Cattani, John Ellis, Jennifer Friedman, Christine Funk, Lynn Garcia, Ted Hunt, David Kaye, Christopher Plourd, and Ronald Reinstein voted in favor of the comments. Dick Reeve and Barry Sheck did not vote. All but one of the current nine members of the committee voted to submit this appeal.

² The full comments are reproduced in Appendix 1 together with the subcommittee’s adjudication as transmitted on June 18, 2018, by Donna Kimball.

³ The Statistics Task Group and NIST scientists expressed much the same concern.

⁴ See, e.g., Daniel J. Capra, Symposium on Forensic Expert Testimony, *Daubert*, and Rule 702, 86 Fordham L. Rev. 1463 (2018), available at: <https://ir.lawnet.fordham.edu/flr/vol86/iss4/1>.

- “If no meaningful differences are noted between the known and unknown samples regarding physical appearance or measurements, then IR spectroscopy should be the next step in the analytical scheme.” § 5.1.
- “When used for spectral comparisons, the objective is to determine whether any meaningful differences exist between the samples.” § 5.3.
- “Spectra cannot be distinguished if they contain no meaningful differences (for example, comparable constituents, reproducible intensities).” § 8.7.

What, then, makes an observed difference “meaningful”? Section 3.1.6 defines a “meaningful difference” as “a feature or property of a sample that does not fall within the variation exhibited by the comparison sample, considering the limitations of the sample or technique, and therefore indicates the two samples do not share a common origin.” A subsection adds that “formal ... statistics” are not used to ascertain “variation exhibited by the comparison sample.”⁵ Neither does the standard specify other methods to help an analyst decide whether observed differences are large enough to indicate that samples do not share a common origin. Therefore, the LRC asked the subcommittee to insert such methods into the standard so that analysts who consult it would have true guidance on “the kinds of variations in presence, positions, shapes and intensities (either in the comparison sample itself or between the two samples being compared) that should be deemed ‘meaningful’” and so that outside observers would know what to look for in spectral comparisons.

Rather than explain why it has not written a standard with meaningful criteria, the subcommittee responded by asserting that “[d]escriptions of how to diagnose if observed differences are meaningful are given in Sections 8.5.1-8.8.” This response cannot be seriously entertained. No real descriptions of how to diagnose whether and when differences in “presence, positions, shapes and intensities” are meaningful are present in these subsections. To show this, we will consider each subsection in turn:

- 8.5.1 There are a number of factors to consider when assessing whether or not spectra can be distinguished from one another: the presence or absence of absorption bands, their positions, shapes, and relative intensities.

This statement does not respond to the LRC comment (and that of various scientists) that there are no criteria for drawing a line between differences in the enumerated factors that are meaningful and those that are not. As the LRC comment noted, “[t]he list of variables to consider (§ 8.5.1) is helpful, for it directs the analyst’s attention to the properties of the spectra that should be compared.” But it fails to tell analysts what to do with measurements of these factors. Merely repeating the words that have been criticized is not an adjudication of the criticism.

- 8.5.1.1 For spectra that cannot be distinguished from one another, characteristic absorption bands observed in one spectrum are also present in the comparison spectrum. The position of the absorption bands should have reasonable agreement with each other and is somewhat dependent on the shape of the absorption band. The positions of corresponding peaks in two

⁵ “Discussion—The use of this term does not imply the formal application of statistics.” § 3.1.6.1. The subcommittee apparently believes that the phrase “exhibited by the comparison sample” refers to the variation seen (or perhaps merely anticipated) in replicate measurements on “the comparison sample.”

or more spectra should be within $\pm 5\text{cm}^{-1}$. Additionally, the absorption bands should have comparable relative intensities and shapes for the spectra being compared.

With the exception of the maximum separation between corresponding peaks,⁶ this subsection is equally devoid of criteria for “meaningful” differences. The words “comparable relative intensities and shapes” and “reasonable agreement” utterly fail to explain what differences would make relative intensities and shapes “comparable” and which ones would constitute “reasonable agreement” for absorption-band positions. Again, merely repeating the words that have been criticized is not an adjudication of the criticism.

- 8.5.1.2 If subtle differences are noted between questioned and known items, where possible, collect additional spectra to demonstrate whether the differences are repeatable and therefore meaningful. The number of additional spectra collected is predicated by several things: the amount of sample present, the hetero-/homogeneity of the material, typical spectral variation observed in similar materials, etc. Therefore, the number of replicates shall be determined on a case-by-case basis.

This subsection empowers analysts to decide that samples are from different sources when the differences are not “subtle.” Criteria for distinguishing between subtle differences and not-so-subtle ones are conspicuously absent. Even if they were present, there are no criteria for deciding how many replicate measurements to make and no indication as to which samples they should be made on: The known sample? The unknown? Both? To say that differences that are “repeatable” are “meaningful” does not fill the gap. Unless there is *no* measurement error, some differences always will recur in a large set of replicate measurements. The subsection is essentially a suggestion that, when an analyst is unsure what to make of differences, he or she should make more measurements and see if that changes his or her opinion. Obviously, repeating a process that has no explicit criteria for deciding what is meaningful does not transform it into a process that has explicit criteria.

- 8.6 Spectra are dissimilar if they contain one or more meaningful differences (for example, absence or presence of constituents, reproducible intensity differences).
- 8.7 Spectra cannot be distinguished if they contain no meaningful differences (for example, comparable constituents, reproducible intensities).

These subsections restate the vacuous definition of “meaningful” in § 5. The parenthetical remarks contain no guidance on how to judge absence or presence of constituents or what makes a difference in intensity (a quantity that is expected to vary, at least a little, in repeated measurements) repeatable.

- 8.8 A spectral comparison is inconclusive if no determination can be made as to whether observed differences are meaningful (for example, peaks are not well resolved, sample condition).

This subsection contains no criteria for reaching conclusions as to what is meaningful. Neither does it contain criteria for choosing not to reach a conclusion. It does not help an analyst

⁶ This quantitative limit seems like a “formal statistical” rule, although the standard does not reveal its origin and validity.

gauge whether peaks are well resolved or whether the sample condition is poor. The standard's treatment of what differences are "meaningful" amounts to a suggestion that analysts should use their best judgment. Hence, the LRC wrote that

It may be that the standard is intended to make these determinations entirely subjective and not governed by any explicit standards. If that is the recommendation or assumption, the document should say so, and it should identify the particular references that estimate the accuracy and reliability of these subjective judgments of "meaningful differences."

The subcommittee's response ignores the substance of this request and cannot be counted as an "adjudication" of it. The subcommittee wrote, first, that

The comment is not persuasive because subjective data evaluations are not necessarily lacking in explicit criteria for assessment. The guidance is explicit that acquired knowledge of the material and the technique must coexist to best evaluate the data, regardless of the evaluation method (visual spectral comparison or mathematical data treatments).

It is difficult to tell what the phrase "acquired knowledge of the material and the technique must coexist" means, but the notion that an injunction to do one's best when evaluating data is an explicit criterion borders on the absurd. Certainly, it is not a response to the observation that under the standard, analysts are expected to make subjective judgments without the benefit of explicit criteria that would promote the reliability of such judgments.

Even so, the LRC did not maintain that subjectivity was necessarily fatal. It asked for nothing more than a standard that would "identify the *particular references* that estimate the accuracy and reliability of these subjective judgments of 'meaningful differences.'" The subcommittee's "adjudication" pointed to no such references. Rather than identify studies that demonstrate the accuracy and reliability of the highly subjective process (under the standard) of spectral comparisons, the subcommittee wrote:

The second sentence is also not persuasive. Known samples can be analyzed and compared using this guide in order to obtain a correct answer (e.g., reference materials compared to a library, proficiency test samples, replicate analyses of a piece of tape). The guide contains citations where the recommended methodology has been used to assess accuracy and reliability.

This paragraph supplies no citation to even a single study that estimates accuracy and reliability. Instead it asserts that it is possible "to obtain a correct answer" with the methods referred to in the standard for collecting data. However, the LRC did not ask whether forensic analysts believe that many of the spectral comparisons of the sort mentioned in the standard are correct. It asked for a scientific estimation of the presumed accuracy and for a similar estimation of the reliability (repeatability and reproducibility) of the subjective analyses. Whether any of the 23 citations in the standard provides this information remains no clearer after the adjudication than it was before the adjudication.

Conclusion

The LRC comments did not question the desirability of rigorous methods for *collecting* data as a part of a larger scheme for making source exclusions and inclusions. The comments were directed at the lack of criteria for *interpreting* spectral data and at the absence of any published validation studies of the interpretative judgments called for by the standard. We recognize that spectral comparisons are not all there is to the analysis of tape components. Our comments only asked that a standard for this part of the process contain explicit criteria for declaring sufficient dissimilarity to exclude the possibility that samples originated from the same source (or for references to studies estimating the accuracy and reliability of largely subjective determinations).

The adjudication does not point to any explicit, articulated criteria for making the judgments called for in the standard. It does not supply references to scientific studies of how analysts making admittedly subjective judgments actually perform. It does not explain why these components of a standard can be dispensed with. It is an adjudication in name only. In substance, it is little more than a conclusory statement of “not persuasive.”

APPENDIX

Adjudication of LRC Comments

(received 6/18/18, as a Word document
with the file name 5c.LRC Comment and Adjudication)



COMMENTS BY THE OSAC LEGAL RESOURCE COMMITTEE (LRC)

TO: Materials (Trace) Subcommittee of the Chemistry-Instrumental SAC
FROM: Lynn Garcia, LRC Liaison to Chemistry-Instrumental Analysis SAC
RE: **OSAC LEGAL RESOURCE COMMITTEE (LRC) COMMENTS ONE3085-17**

Introduction

The Chemistry SAC has proposed the addition of the recently revised “Standard Guide for Fourier Transform Infrared Spectroscopy in Forensic Tape Examinations” (ASTM E3085-17) to the OSAC Registry of Approved Standards. On the basis of the information that has been provided to us, we believe that, with certain modifications, the standard can be included in the Registry of Approved Standards. Absent these modifications however, the LRC does not believe the standard guide is ready for the Registry.

Guidelines on Ascertaining “Meaningful Differences” Are Missing

The standard defines a “meaningful difference” as “a feature or property of a sample that does not fall within the variation exhibited by the comparison sample, considering the limitations of the sample or technique, and therefore indicates the two samples do not share a common origin.” § 3.1.6. The standard adds that “[t]he use of this term does not imply the formal application of statistics.” § 3.1.6.1.

Our major concern is that the remainder of the standard offers no guidance on how to ascertain whether a difference is meaningful.⁷The standard states “Spectra cannot be distinguished if they contain no meaningful differences,” and then some examples are given.

The list of variables to consider (§ 8.5.1) is helpful, for it directs the analyst’s attention to the properties of the spectra that should be compared. But it is not clear what “variation exhibited by the comparison sample”

⁷We also note that the phrasing of § 3.1.6.1 is awkward—the use of the term “meaningful difference” does not imply the informal application of explicit statistics either, although it necessarily rests on statistical reasoning about variability. We suspect the sentence should be rephrased as follows: “Meaningful differences can be discerned without a quantitative analysis.” This observation is an editorial comment. If the sentence on statistical reasoning were the only item to be revised, it would not prevent approval of the standard.

(§ 3.1.6) means for these properties. Is there any variation exhibited by the comparison sample in “the presence or absence of absorption bands, their positions, shapes, and relative intensities”? Isn’t a band either present or absent? Whichever it is, there is no variation. Likewise, the position is where it is. The shape and relative intensity are what they are. The subsections do not describe the kinds of variations in presence, positions, shapes and intensities (either in the comparison sample itself or between the two samples being compared) that should be deemed “meaningful.”

It may be that the standard is intended to make these determinations entirely subjective and not governed by any explicit standards. If that is the recommendation or assumption, the document should say so, and it should identify the particular references that estimate the accuracy and reliability of these subjective judgments of “meaningful differences.”

Documentation

We applaud the inclusion of Section 9, on documentation, in the standard. Consideration should be given to including a further section on reporting and testifying. If these matters are to be addressed in a separate standard—such as the *Standard Practice for Interpretation and Report Writing in Forensic Comparison of Trace Materials*—that fact should be mentioned.

Section 9.1 states that “[w]hen making comparisons of tape samples, the analyst’s assessment of the IR spectra shall be documented.” To provide guidance on what is required or recommended, this section should describe the type of information that should be in this documentation. For example, the analyst might be asked to annotate two spectra to show the differences that were considered.

Section 9.3 states that “[c]ase notes shall include a copy of the instrumental data that was used to reach a conclusion. All paper and electronic copies that are retained as part of the case file shall include a unique sample designation, the operator’s name or initials, and the date of analysis.” The word “includes” connotes that this list does not exhaust the necessary contents of the case notes, and one can infer from Section 9.1 that the documentation of the assessment of the spectra must be in the case notes, but this should be made explicit by stating that the case notes should also include the documentation of the assessment made by analyst on the spectra as well as all information necessary for an independent examiner to conduct a later analysis of the original analysis and the conclusions reached.

Section 9.4 states that “[a] description of the evidence analyzed by IR, the method of sample preparation, the analytical instrumentation used, mode of operation (transmission, ATR, etc.), and its operating parameters shall be included in the case notes or in the procedural manuals.” The case notes themselves should contain the operating parameters. At a minimum, they must include a notation that information is available in the instrument manuals.

DISCLAIMER: *The failure of any member of the Legal Resource committee (LRC) to provide a comment, identify a legal issue or join in another LRC comment should not be interpreted as a disagreement or endorsement of the comment, the standard or its legal sufficiency.*