Designation: E2214 – 20

Standard Practice for Specifying and Verifying the Performance of Color-Measuring Instruments[1]

This standard is issued under the fixed designation E2214; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

INTRODUCTION

Recent advances in optics, electronics and documentary standard have resulted in a proliferation of instruments for the measurement of color and appearance of materials and objects. These instruments possess very good performance but there has been little progress toward standardizing the terminology and procedures to quantify that performance. Therefore, the commercial literature and even some documentary standards are a mass of confusing terms, numbers and specifications that are impossible to compare or interpret.

Two recent papers in the literature, have proposed terms and procedures to standardize the specification, comparison and verification of the level of performance of a color-measuring instrument.[2,3] Following those procedures, those specifications can be compared to product tolerances. This becomes important so that instrument users and instrument makers can agree on how to compare or verify, or both, that their instruments are performing in the field as they were designed and tested in the factory.

1. Scope

1.1 This practice provides standard terms and procedures for describing and characterizing the performance of spectral and filter based instruments designed to measure and compute the colorimetric properties of materials and objects. It does not set the specifications but rather gives the format and process by which specifications can be determined, communicated and verified.

1.2 This practice does not describe methods that are generally applicable to visible-range spectroscopic instruments used for analytical chemistry (UV-VIS spectrophotometers). ASTM Committee E13 on Molecular Spectroscopy and Chromatography includes such procedures in standards under their jurisdiction.

1.3 *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, health, and environmental practices and determine the applicability of regulatory limitations prior to use.*

1.4 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

2. Referenced Documents

2.1 *ASTM Standards:*[4]

D2244 Practice for Calculation of Color Tolerances and Color Differences from Instrumentally Measured Color Coordinates

E284 Terminology of Appearance

E1164 Practice for Obtaining Spectrometric Data for Object-Color Evaluation

[1] This practice is under the jurisdiction of ASTM Committee E12 on Color and Appearance and is the direct responsibility of Subcommittee E12.04 on Color and Appearance Analysis.

Current edition approved May 1, 2020. Published May 2020. Originally approved in 2002. Last previous edition approved in 2019 as E2214 – 19. DOI: 10.1520/E2214-20.

[2] Ladson, J., "Colorimetric Data Comparison of Bench-Top and Portable Instruments," *AIC Interim Meeting, Colorimetry*, Berlin, 1995.

[3] Rich, D., "Standardized Terminology and Procedures for Specifying and Verifying the Performance of Spectrocolorimeters," *AIC Color 97 Kyoto*, Kyoto, 1997.

[4] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

2.2 *Other Documents:*

ISO VIM International Vocabulary of Basic and General Terms in Metrology (VIM)[5]

NIST Technical Note 1297 Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results[6]

3. Terminology

3.1 Definitions of appearance terms in Terminology E284 are applicable to this practice.

3.2 Definitions of metrology terms in ISO, International Vocabulary of Basic and General Terms in Metrology (VIM) are applicable to this practice.

3.3 *Definitions of Terms Specific to This Standard:*

3.3.1 *colorimetric spectrometer, n*—spectrometer, one component of which is a dispersive element (such as a prism, grating or interference filter or wedge or tunable or discrete series of monochromatic sources), that is normally capable of producing as output colorimetric data (such as tristimulus values and derived color coordinates or indices of appearance attributes) as well as the underlying spectral data from which the colorimetric data are derived.

3.3.2 *inter-instrument agreement, n*—the closeness of agreement between the results of measurements in which two or more instruments from the same manufacturer and model are compared.

3.3.3 *inter-model agreement, n*—the closeness of agreement between the results of measurements in which two or more instruments from different manufacturers, or of different but equivalent design, are compared.

3.3.3.1 *Discussion*—Modern instruments have such high precision that small differences in geometric and spectral design can result in significant differences in the performance of two instruments. This can occur even though both instruments exhibit design and performance bias which are well within the expected combined uncertainty of the instrument and within the requirements of any international standard.

4. Summary of Practice

4.1 This practice defines standardized terms for the most common instrument measurement performance parameters (repeatability, reproducibility, inter-instrument agreement, inter-model instrument agreement, accuracy) and describes a set of measurements and artifacts, with which both the producers and users of color-measuring instruments verify or certify the specification and performance of color-measuring instruments. Following this practice can improve communications between instrument manufacturers and instrument users and between suppliers and purchasers of colored materials.

[5] ISO/IDE/OIML/BIPM, *International Vocabulary of Basic and General Terms in Metrology*, International Organization for Standardization, Geneva, Switzerland, 1984.

[6] Taylor, Barry N., and Kuyatt, Chris E., *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297, U. S. Government Printing Office, Washington, D. C., 1984.

5. Significance and Use

5.1 In today's commerce, instrument makers and instrument users must deal with a large array of bench-top and portable color-measuring instruments, many with different geometric and spectral characteristics. At the same time, manufacturers of colored goods are adopting quality management systems that require periodic verification of the performance of the instruments that are critical to the quality of the final product. The technology involved in optics and electro-optics has progressed greatly over the last decade. The result has been a generation of instruments that are both more affordable and higher in performance. What had been a tool for the research laboratory is now available to the retail point of sale, to manufacturing, to design and to corporate communications. New documentary standards have been published that encourage the use of colorimeters, spectrocolorimeters, and colorimetric spetrometers in applications previously dominated by visual expertise or by filter densitometers.[7] Therefore, it is necessary to determine if an instrument is suitable to the application and to verify that an instrument or instruments are working within the required operating parameters.

5.2 This practice provides descriptions of some common instrumental parameters that relate to the way an instrument will contribute to the quality and consistency of the production of colored goods. It also describes some of the material standards required to assess the performance of a color-measuring instrument and suggests some tests and test reports to aid in verifying the performance of the instrument relative to its intended application.

6. Instrument Performance Parameters

6.1 *Repeatability* is generally the most important specification in a color-measuring instrument. Colorimetry is primarily a relative or differential measurement, not an absolute measurement. In colorimetry, there is always a standard and a trial specimen. The standard may be a real physical specimen or it may be a set of theoretical target values. The trial is usually similar to the standard in both appearance and spectral nature. Thus, industrial colorimetry is generally a test of how well the instrument repeats its readings of the same or nearly the same specimen over a period of minutes, hours, days, and weeks.

6.1.1 The ISO VIM defines repeatability as a measure of the random error of a reading and assumes that the sample standard deviation is an estimate of repeatability. Repeatability is further defined as the standard deviation of a set of measurements taken over a specified time period by a single operator, on a single instrument with a single specimen. This definition is similar to that in Terminology E284, except that the ISO explicitly defines the metric of "closeness of agreement" as the sample standard deviation. Since color is a multidimensional property of a material, repeatability should be reported in terms of the multidimensional variance–covariance matrix, or in terms of the 95 % confidence interval of its combinatorial color difference for a single specimen. See 6.6.

[7] ISO 13655 Spectral Measurement and Colorimetric Computation for Graphic Arts Images, International Organization for Standardization, Geneva, Switzerland.

6.1.2 The time period over which the readings are collected must be specified and is often qualitatively described as "short," "medium," or "long." The definitions of these time frames do not overlap. This is intentional, providing clearly defined milestones in the temporal stability of test results.

6.1.2.1 For the purposes of colorimetry, "short" is normally the time required to collect a set of 30 readings, taken as fast as the instrument will allow. The actual time will vary as a function of lamp and power supply characteristics but should be less than one hour.

6.1.2.2 "Medium" term is normally defined as, at least the period of one work shift (8 h) but less than three work shifts (one day).

6.1.2.3 "Long" term is open ended but is often described as any set readings taken over a period of at least 4 to 8 weeks. The longest known reported study described readings taken over a period of 3¼ years.[8]

6.2 *Reproducibility* is the second most important specification in a color-measuring instrument. According to Terminology E284, reproducibility is the closeness of agreement of the results of measurements in which one or more of the measurement parameters have been systematically changed. Thus the sample is different, the procedures or instrument are different, or the time frame is very long. The increase of disorder over a very long time changes the instrument systematically and the set of readings really compares a "young" instrument with an "old" instrument.

6.2.1 The ISO VIM defines reproducibility as the closeness of agreement of the results of measurements in which either the time frame is very long, in which the operator changes, the instrument changes, or the measurement conditions change. ISO again recommends estimating this with a standard deviation. Reproducibility is further defined as the standard deviation of a set of measurements taken over a specified period of time by a single operator, on a single instrument with a single specimen. This definition is similar to that in Terminology E284, except that the ISO again, explicitly defines the metric of "closeness of agreement" as the sample standard deviation. Again, since color is a multidimensional property of a material, reproducibility should be reported in terms of the multidimensional variance–covariance matrix.

6.2.2 The time period over which the readings are collected must be specified. Repeatability and reproducibility have traditionally been evaluated in colorimetry by comparing the color differences of a set of readings to a single reading or to the average of the set of readings.

6.3 *Inter-Instrument Agreement,* as defined in 3.3.2, describes the reproducibility between two or more instruments, of identical design. The ISO has no definition or description of such a concept. This is because in most test results, a method or instrument dependent bias can be assessed. In this situation, such a test measures the consistency of the design and manufacturing process. Within the technical description of the standard geometric and spectral parameters for the measurement of diffuse reflectance factor and color, a significant amount of latitude exists. This latitude results in a random amount of bias. For a given design, a manufacturer may reduce the random bias, often to a level less than the stability of reference materials. The most common form of test for inter-model instrument agreement is pairwise color difference assessment of a series of specimens. Various parameters are reported in the literature including the average color difference, the maximum color difference, the typical color difference, the RMS color difference or the MCDM mean color difference from the mean, taking the average of all instruments as the standard and the other as the test instrument. Using pairs of instruments and materials one can derive a multivariate confidence interval against the value 0.0 difference and then test individual components to determine which attribute (lightness, chroma, hue) are the significant contributors to the differences between instruments. If a group of instruments are being tested then a multivariate analysis of variance (MANOVA) can be performed to test the agreement of the means of the instrument.

6.4 *Inter-Model Agreement,* as defined in 3.3.3, describes the reproducibility between two or more instruments of differing design. The latitude within the standard geometric and spectral parameters described in the preceding paragraph is at a maximum when the designs differ. The systematic bias may increase by factors of from 5 to 10 because of the increased latitude. Standardizing laboratories will report either the algebraic differences between measurement results or the ratio of the measurement values between two labs. The former will be a normal statistical variable if the measurement values are normally distributed, and the latter will be distributed as a ratio of normally distributed variables. This distribution can be estimated from the multivariate variance–covariance matrix. Using pairs of instruments and materials one can derive a multivariate confidence interval against the value 0.0 difference and then test individual components to determine which attribute (lightness, chroma, hue) are the significant contributors to the differences between instruments. If a group of instruments are being tested then a multivariate analysis of variance (MANOVA) can be performed to test the agreement of the means of the instrument.

6.5 *Accuracy,* while occasionally critical, is generally the least significant parameter in characterizing the performance of a color-measuring instrument. ISO defines accuracy as the conformance of a series of readings to the accepted or true value. In modern colorimetry, the volume of the total combined uncertainty around the accepted value is often many times larger than volume of visual acceptability of the products whose color is being quantified. Therefore, an "accurate" color measurement may result in an unacceptable product color. There are two scales in a spectrocolorimeter that can be assigned nominal values and tested against standard values. They are the radiometric scale and the wavelength scale.

6.5.1 The wavelength scale includes the sampling position (centroid wavelength) and the sampling window width (spectral bandwidth). These parameters are normally tested against physical standards of wavelength based on fundamental phenomena, such as discharge lamps or laser lines. In very

[8] Rich, D. C., Battle, D., Malkin, F., Williamson, C., Ingleson, A., "Evaluation of the Long-Term Repeatability of Reflectance Spectrophotometers," *Spectrophotometry, Luminescence and Colour: Science and Compliance*, C. Burgess and D. G. Jones, eds., Elsevier, Amsterdam, 1995.