

Goniochromatic assessment of gray scales for color change

ALEJANDRO FERRERO,^{1,*} BERTA BERNAD,¹ JOAQUÍN CAMPOS,¹ NOËL RICHARD,² CHRISTINE FERNÁNDEZ-MALOIGNE,² AND MANUEL MELGOSA³

¹ *Instituto de Óptica “Daza de Valdés”, Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain*

² *XLIM Laboratory, UMR 7252 CNRS, University of Poitiers, Poitiers, France*

³ *Department of Optics, Faculty of Sciences, University of Granada, Granada, Spain*

[*alejandro.ferrero@csic.es](mailto:alejandro.ferrero@csic.es)

Abstract: The dependence of color differences on the illumination and viewing directions for two widely-used gray scales for color change (SDCE and AATCC) was evaluated through measuring the spectral bidirectional reflectance distribution function (BRDF) by a gonio-spectrophotometer of metrological quality. Large incidence and viewing angles must be specially avoided using these gray scales because in these conditions color differences vary considerably from those established in ISO 105-A02 and ASTM D2616-12. While the visual appearance of the SDCE and AATCC gray scales for color change is similar, our results indicate that their goniochromatic properties are very different. Finally, some recommendations to use these gray scales for visual experiments are given.

OCIS codes: (330.1690) Color; (330.1710) Color, measurement; (330.1730) Colorimetry.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Color fastness is a term used in the dyeing of textile materials, meaning resistance of the material's color to fading or running when subjected to a particular treatment or environment [1]. Specific gray scales were developed long ago [2] and since then have been used in the textile industry for assessing color changes in fastness tests. The standard procedure is performed by an experienced operator, under specific illumination and viewing conditions, who visually compares the difference of two samples with respect to the pairs in a gray scale, and provides the magnitude for the given color change [3-5]. In a similar way, gray scales have been also used as reference in many visual experiments to evaluate the magnitude of color and appearance differences in sample pairs [6-13].

Gray scales for color change were designed to be used under the CIE bidirectional geometry 45°:0° [14] (angle of incidence of 45° with respect to the normal, and frontal viewing). Their use at other geometries might have an impact in the fastness tests and visual assessment results, in particular if there is a color dependence on the geometry due to surface treatments. In practice, the CIE bidirectional geometry 45°:0° is not perfectly realized. One of the reasons is that the illumination is not completely directional, since the luminous sources under which the pairs are observed are not point sources, but extended sources. For example, gray scales for color change have been used in multi-angle color-assessment cabinets with large fluorescent tubes [11, 12], and, consequently, with illumination consisting of a broad bunch of different directions, from one extreme of the tube to the other. In addition, due to geometrical restrictions in color-assessment cabinets, the viewing angle cannot be exactly 0° for the evaluated pair and the pairs of the gray scale. In a previous study [15], we noticed an important dependence of the performance of a gray scale for color change on the observation geometry when used in a multi-angle color-assessment cabinet. Accurate results in visual

experiments require accurate spectrophotometric measurements of reference materials, such as reflectance standards or gray scales [16].

In this study, two recently acquired specimens of commercial gray scales for color change from two different manufacturers were measured using a gonio-spectrophotometer of well-contrasted metrological quality to make accurate measurements under bidirectional conditions [17]. Both commercial scales, called “SDCE Grey Scale for Assessing Change in Colour (Society of Dyers and Colourists Enterprises Limited)” and “AATCC Gray Scale for Evaluating Change in Color (American Association of Textile Chemists and Colorists)”, hereafter referred respectively as SDCE and AATCC gray scales, are similarly arranged (see Fig. 1). They consist of nine color pairs of non-glossy neutral gray chips of identical dimensions, which realize color differences with increasing magnitudes from 0 to about 15 CIELAB units.

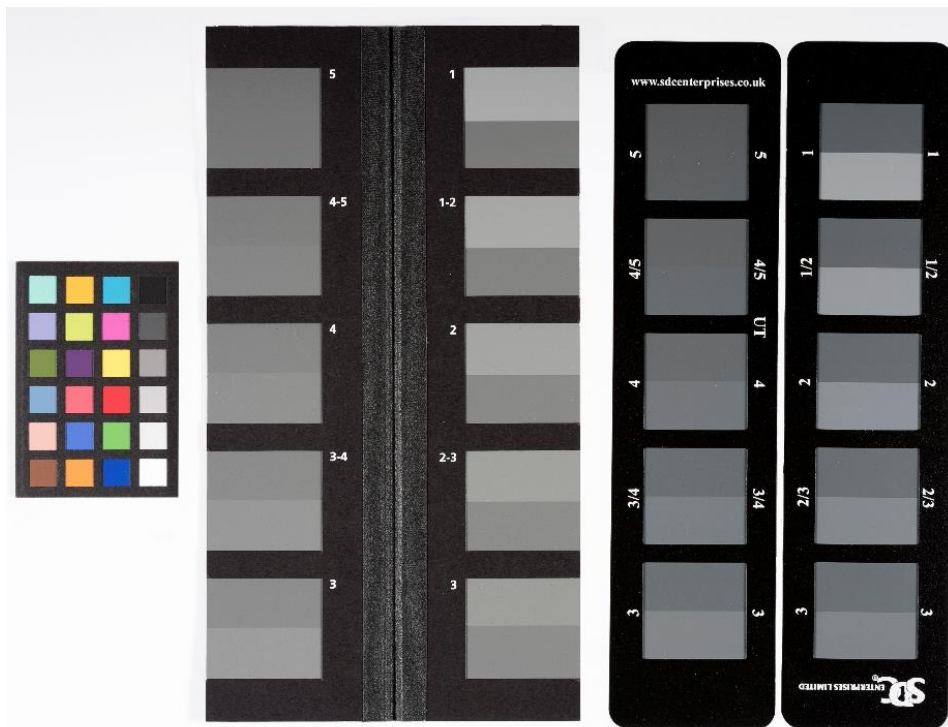


Fig. 1: Picture of the pairs in the AATCC (left) and SDCE (right) gray scales for color change. The pairs in the SDCE scale are in the front and back of one strip. Note that grade 3 color pair is repeated in both gray scales (with a top-bottom change of the chips in the AATCC scale), and grade 5 is a color pair with two physically identical chips.

The standardized values of color differences in CIELAB units [14], ΔE^*_{ab} , are shown in Table 1 for the different grades, G (different pairs in the gray scale for color change), together with their corresponding tolerances ($\delta[\Delta E^*_{ab}]$, after the symbol \pm). The CIEDE2000 color-difference formula, currently recommended by CIE and ISO for pairs with color differences below 5.0 CIELAB units [18], is not applicable in this work because of the larger differences presented by the gray scale’s lower grades (Table 1). To better realize the visual difference of these pairs, it is noticed that a CIELAB unity is close to the just-noticeable visual difference for observers with normal color vision [19-20].

Table 1: CIELAB color differences \pm tolerances for the nine pairs of the standard gray scale for color change, as specified by ISO 105-A02 and ASTM D2616-12.

Grade (G)	CIELAB color difference $\Delta E_{ab}^* \pm \delta[\Delta E_{ab}^*]$
5	0.0 ± 0.2
4-5	0.8 ± 0.2
4	1.7 ± 0.3
3-4	2.5 ± 0.3
3	3.4 ± 0.4
2-3	4.8 ± 0.5
2	6.8 ± 0.6
1-2	9.6 ± 0.7
1	13.6 ± 1.0

The color pairs in Table 1 provide ratings of 5 (no color contrast), 4-5, 4, 3-4, 3, 2-3, 2, 1-2 and 1 (highest color contrast). The ratings 4-5, 3-4, 2-3, and 1-2 are sometimes designated as 4.5, 3.5, 2.5, and 1.5, respectively, but the former nomenclature is used in this paper. Grade 5 is represented on the real gray scales by two identical adjacent chips. Grades from 4-5 to 1, both included, are represented by chips with the same lightness of those in grade 5 paired with increasingly lighter chips (Fig. 1). Note that the gray scales here studied show a geometric increase of color differences, although gray scales with linear increase have been suggested for visual tests of small color differences [21-22]. In this paper, reflectance and color difference measurements of the chips in the SDCE and AATCC gray scales are reported. These measurements were carried out under very well defined illumination and viewing angles (bidirectional geometries), allowing a physical characterization of the angular dependence to be achieved.

2. Experimental methods

In order to evaluate the dependence on illumination and viewing directions of the color of each neutral chip, the spectral bidirectional reflectance distribution function (BRDF) at a well-distributed set of geometries was measured; that is, at different pairs of irradiation and detection directions uniformly covering the incidence plane. The spectral BRDF describes the reflection properties of elementary surfaces as a function of specific irradiation and detection directions [23-24], and it is directly related to the spectral reflectance factor R at the same geometrical conditions, as $R = \pi \times \text{BRDF}$, which is required to calculate primary color coordinates such as tristimulus values [14].

The spectral BRDF measurements were performed using the gonio-spectrophotometer GEFE, an instrument designed and developed at *Instituto de Óptica “Daza de Valdés” (Consejo Superior de Investigaciones Científicas, Madrid)*. It was conceived to measure the spectral BRDF at any pair of irradiation and detection directions, including out-of-plane and retro-reflection geometries [17]. It comprises three systems: the irradiation, the sample-positioning system, and the detection system. The first one is fixed, whereas the other two systems are mobile: the sample is placed with the required orientation relative to the incoming beam, while the detector is attached to a cogwheel to be able to revolve around the sample, allowing a fast and accurate sampling. GEFE operates with a wide-band xenon lamp, which emits in the 185 nm - 2000 nm spectral range. To irradiate the samples uniformly with a

highly directional beam, a Köhler optical system is used. The detection is carried out by a spectroradiometer (CS-2000 A, Konica Minolta) that relies on a CCD detector. Its spectral response includes the visible range (from 380 nm to 780 nm) and it has a variable field of view (0.1°, 0.2°, and 1°), although only the field of view of 1° has been used in the current work. Finally, a beam-splitter allows retro-reflection measurements.

For each of the 9 different chips in both gray scales, the CIELAB color coordinates at every geometry were calculated from the spectral BRDF measurements, using the CIE standard illuminant D65 and CIE 1964 standard colorimetric observer [14]. The uncertainty of the color differences in the color pairs of the gray scales was estimated by a Monte Carlo procedure (1000 repetitions), by using the BRDFs and their uncertainties to simulate a pool of possible BRDFs and calculate the standard deviation of the color differences. This standard deviation was consistently below 0.04 CIELAB units at 95% confidence level, which may be considered a negligible value in comparison with human just-noticeable visual differences [19-20].

In the following, the irradiation and detection directions are denoted by θ , which is the angle with respect to the normal direction to the chip surface. Irradiation is denoted by the subscript 'i' and detection by the subscript 's' (from scattering). Every geometry consists of an irradiation direction and a detection direction. Six different irradiation directions were selected for this study, from $\theta_i = 0^\circ$ to 75° , with steps of 15° . For each of these directions, the spectral BRDF was measured for 11 detection directions, from $\theta_s = -75^\circ$ to 75° , with steps of 15° too, where negative values mean that the detection direction is within the same incidence half-plane containing the irradiation direction. For instance, $\theta_i = 45^\circ$ and $\theta_s = 45^\circ$ denotes a geometry to measure specular reflection, whereas $\theta_i = 45^\circ$ and $\theta_s = -45^\circ$ denotes a retro-reflection condition. In this paper, a given measurement geometry is expressed as $\theta_i:\theta_s$.

3. Results and discussion

3.1 Spectrally-averaged BRDF

No important wavelength dependence of the BRDF is observed in the measurements, as expected from achromatic surfaces. The measured spectrally-averaged BRDF corresponding to the 9 chips in the AATCC and SDCE gray scales are shown in Fig. 2 and Fig. 3, respectively. They are the averaged values of two available different specimens of the above-mentioned gray scales. It must be noticed that each chip is denoted here with the grade of the pair where it is located (chip 5 is the darkest chip, the common chip to all pairs, and also the duplicated chip in pair with grade 5, which has a null color difference). For example, the chip 2-3 is the lightest chips in the color pair (grade) 2-3.

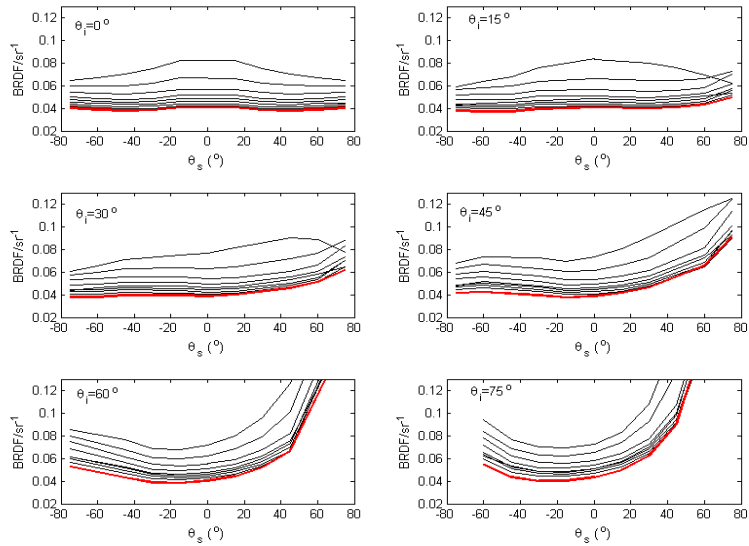


Fig. 2: Spectrally-averaged BRDF measurements corresponding to the 9 chips of the AATCC gray scale at six different irradiation angles (average of the two specimens). For easier comparison these six subplots have the same y-axis scale (the BRDF values found for $\theta_i = 75^\circ$ and the highest θ_s have been missed). The red curves correspond to results for chip 5 (the darkest one).

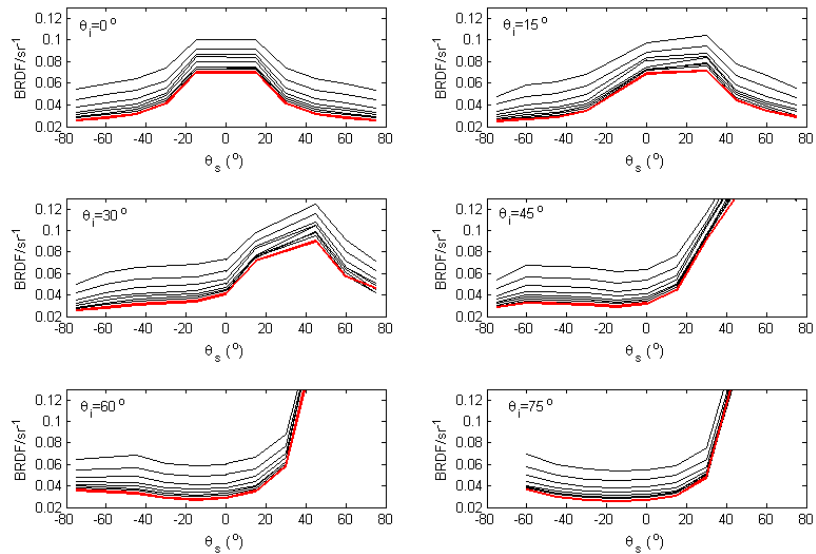


Fig. 3: Same as in Fig. 2, but for the chips of the SDCE gray scale.

Each of the 6 subplots in Fig. 2 and Fig. 3 represents the spectrally-averaged BRDF at a specific angle of incidence (0° , 15° , 30° , 45° , 60° , and 75°) as a function of the detection angle. Each curve in these subplots corresponds to a different chip, the red thickest curve representing the result for chip 5, which is the darkest chip in the gray scales. Since the lightness of the chips is correlated with the BRDF, it is simple to assign the correct chip number to every curve (from 5 to 1) in the subplots. The expanded relative uncertainty of every measure at a given wavelength and geometry was estimated to be around 1 %, and it accounts for uncertainty contributions from the comparison with the standard, repeatability, reproducibility, linearity of the detector, uniformity of the irradiation on the chip, and angular positioning.

In absence of any goniochromatic effect (i.e. perfect diffuse reflectance), the curves in Fig. 2 and Fig. 3 should be horizontal lines with different y-axis values (the highest one for chip 1 and the lowest one for chip 5). However, this is not the case in Fig. 2 and Fig. 3, where it seems that most perfect diffuse chips are those in the AATCC scale under the lowest incidence angles. For both gray scales, a lower symmetry with respect to the specular geometry $\theta_s = \theta_i$ is found for higher incidence angles. The steady increase of the BRDF for very high incidence and reflection angles is typical of diffuse surfaces [25], and it is related to the appearance of sheen, or shininess at grazing angles in otherwise matte specimens [26]. This increase is higher in the case of the chips in the SDCE gray scale. In addition, the largest increase of the BRDF around the specular direction is found for the SDCE scale. In general, the dependence on the geometry of the spectrally-averaged BRDF is lower for the chips in the AATCC gray scale (Fig. 2) than for the ones in the SDCE gray scale (Fig. 3). These results are coherent with gloss measurements performed using a Multi-Gloss 268 (Konica Minolta Sensing). The gloss of the chips in the SDCE gray scale was in the range 4.4 GU - 5.4 GU (average 4.9 GU) for 60° , and 9.5 GU - 12.1 GU (average 10.8 GU) for 85° ; and the gloss of the chips in the AATCC gray scale was in the range 0.6 GU - 1.6 GU (average 0.8 GU) for 60° and 1.5 GU - 2.8 GU (average 1.8 GU) for 85° . Therefore, although both gray scales are very matte, the average gloss of the SDCE scale is around 6 times higher than AATCC scale's. Finally, for two chips with identical designation, the values of the spectral BRDFs in the SDCE gray scale are generally higher than in the AATCC gray scale.

The spectral BRDF data shown in Fig. 2 and Fig. 3 are available in Dataset 1 (Ref. [;Error! No se encuentra el origen de la referencia.]).

3.2 Colorimetric analyses of individual chips

From BRDF measurements, assuming the CIE standard illuminant D65 and CIE 1964 standard colorimetric observer, tristimulus values and CIELAB color coordinates of the chips in the gray scales were calculated for each geometry [14].

Because two recently acquired specimens of both gray scales were measured, the consistency of chips produced by each manufacturer could be analyzed from CIELAB color differences between two chips with identical designation at each measurement geometry. The results of this analysis for the AATCC and SDCE gray scales are shown in Table 2, where cells show average CIELAB color difference for the 9 chips. Last column in Table 2 shows that color differences increase with the incidence angle for both gray scales, values for AATCC being higher than for SDCE. Highest values in Table 2 correspond to the retro-reflection and specular detection directions. Excluding these two directions, average color differences in Table 2 are small (maximum values of 1.4 and 0.7 CIELAB units for AATCC and SDCE gray scales, respectively). Under recommended standard conditions ($\theta_i = 45^\circ$; $\theta_s = 0^\circ$) [3-5], Table 2 shows the average color difference between chips from the same manufacturer was 0.5 and 0.2 CIELAB units for AATCC and SCDE, respectively. Since the just noticeable visual difference is around 1.0 CIELAB unit, the results in Table 2 indicate that the two measured specimens in both gray scales are visually undistinguishable, except for

largest incidence and reflection angles. Because of this reason, in this paper average results of two measured specimens from each manufacturer are reported.

Table 2. Average CIELAB color differences between two chips with identical designations in two new gray scales manufactured by AATCC and SDCE. The results in each cell correspond to the average of the 9 chips for each incidence angle θ_i (first column) and each reflection angle θ_s . Last column shows average values for each incidence angle (i.e. average of all reflection angles).

AATCC												
θ_i / θ_s	-75°	-60°	-45°	-30°	-15°	0°	15°	30°	45°	60°	75°	Ave.
0°	0.2	0.2	0.2	0.3	0.3	0.5	0.3	0.3	0.2	0.2	0.2	0.3
15°	0.8	0.3	0.2	0.3	0.7	0.4	0.4	0.3	0.4	0.6	3.6	0.7
30°	0.8	0.3	0.2	1.0	0.3	0.4	0.5	0.4	0.4	0.4	4.4	0.8
45°	0.6	0.3	3.8	0.3	0.5	0.5	0.6	0.6	0.7	1.0	3.9	1.2
60°	0.7	13.6	0.4	0.5	0.6	0.6	0.7	0.8	1.3	1.9	3.1	2.2
75°	10.5	1.2	0.8	0.9	0.9	1.0	1.0	1.5	2.2	3.3	9.4	3.0

SDCE												
θ_i / θ_s	-75°	-60°	-45°	-30°	-15°	0°	15°	30°	45°	60°	75°	Ave.
0°	0.1	0.1	0.1	0.3	0.7	0.9	0.8	0.3	0.1	0.1	0.1	0.3
15°	0.4	0.1	0.1	0.2	0.5	0.7	0.9	0.5	0.3	0.2	0.5	0.4
30°	0.3	0.1	0.1	0.9	0.1	0.3	0.8	1.1	0.5	0.4	0.8	0.5
45°	0.3	0.1	1.5	0.1	0.2	0.2	0.3	1.0	1.6	0.7	1.2	0.6
60°	0.4	1.4	0.2	0.2	0.2	0.2	0.2	0.4	1.2	2.5	1.8	0.8
75°	2.2	0.9	0.4	0.4	0.3	0.3	0.3	0.4	0.9	2.5	9.3	1.6

CIELAB lightness differences between chips with the same designations in the SDCE and AATCC gray scales are shown in Table 3, for two incidence angles (15° and 45°, which are the most common ones in multi-angle measurements) and six reflection angles. The values before/after \pm in Table 3 represent the average/standard deviations across the 9 chips with different lightness. The results in Table 3 show that the chips in the SDCE gray scale are considerably lighter than those in the AATCC gray scale around the specular angles (15°:0°, 15°:30°, 45°:30°, and 45°:60°), the opposite being true for large aspecular angles.

Table 3: CIELAB lightness differences between chips with identical designations in the two gray scales, at incidence angles of 15° and 45°, considering 6 reflection angles. The values before/after \pm represent the average/standard deviations across the 9 chips with different lightness.

Lightness differences	
Bidirectional	$L_{10,SDCE}^i - L_{10,AATCC}^i$
15°:-75°	-7.4 \pm 1.0
15°:0°	9.2 \pm 2.6
15°:30°	10.8 \pm 2.2
15°:45°	0.7 \pm 0.5
15°:60°	-4.3 \pm 0.7
15°:75°	-10.8 \pm 3.7
45°:-75°	-7.6 \pm 1.0
45°:0°	-3.6 \pm 0.4
45°:15°	0.5 \pm 0.8

45°:30°	13.6 ± 3.3
45°:60°	25.1 ± 4.1
45°:75°	8.7 ± 2.3

The results in Table 3 are consistent with the spectral reflectance factors shown in Fig. 4, where, for simplicity, only three of the nine chips in the gray scales have been selected, at the same geometries given in Table 3 for $\theta_i = 45^\circ$. For the theoretical case of an achromatic perfect diffuser, spectral reflectance factors should be independent of the wavelength. This is not the case in Fig. 4, with a larger spectral variation at the more grazing viewing angles and for wavelengths lower than around 450 nm, where there is a notable spectral feature, which could be a combined effect of polarizing properties of the surface and some polarization sensitivity of the measuring system. It also can be observed in Fig. 4 that the variability with the geometry is higher for SDCE (right column) than for AATCC (left column). In particular, it can be noted in Fig. 4 that the spectral reflectance factors are higher for the chips in the SDCE for geometries 45°:30°, 45°:60° and 45°:75°, which led to higher lightness values of chips of the SDCE scale in these three geometries (Table 3).

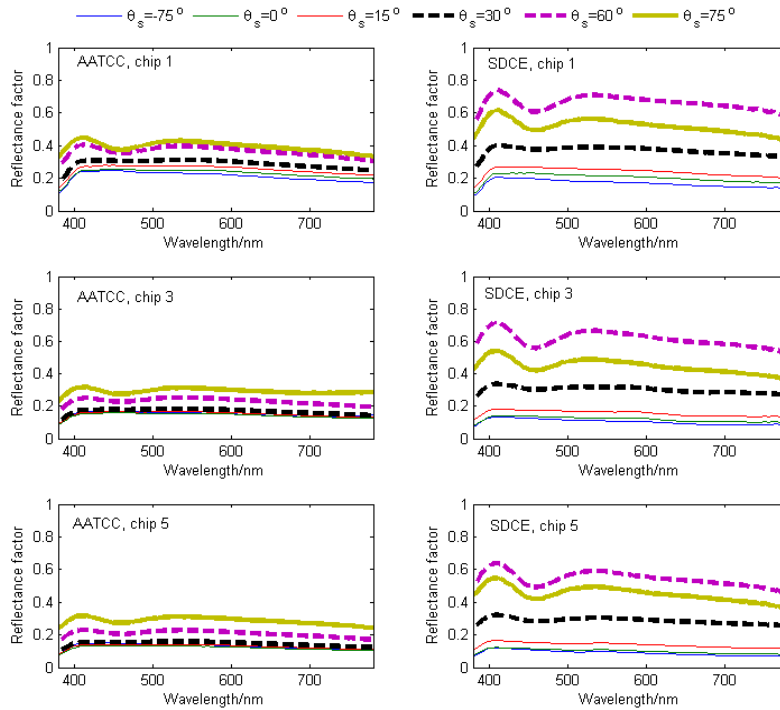


Fig. 4: Spectral reflectance factor for 3 of the 9 chips in the AATCC (left) and SDCE (right) gray scales, under six of the geometries shown in Table 3 with $\theta_i = 45^\circ$ (45°:75°, 45°:0°, 45°:15°, 45°:30°, 45°:60°, and 45°:75°).

3.3 Colorimetric analyses of color pairs

The gray scales for color change can no longer be used at those geometries for which color differences of their color pairs are outside the tolerances specified in Table 1 [3-5]. A parameter $K(G_i)$ was calculated for the studied geometries as:

$$K(G_i) = \frac{|\Delta E_{ab,STD}^*(G_i) - \Delta E_{ab,MEA}^*(G_i)|}{\delta[\Delta E_{ab,STD}^*(G_i)]} \quad (1)$$

where $\Delta E_{ab,MEA}^*(G_i)$ is the measured CIELAB color difference between the two chips at grade i (G_i), and $\Delta E_{ab,STD}^*(G_i)$ and $\delta[\Delta E_{ab,STD}^*(G_i)]$ are the CIELAB color differences and corresponding tolerances for the same grade i , as given in Refs. [3-5] (Table 1). Note that, whereas only individual chips in gray scales have been previously discussed (subsections 3.1 and 3.2), now the analysis is focused on the color differences of the nine pairs (grades) in the gray scales, which is the most important concept in usual practice.

If $K(G_i)$ is lower than 1, the pair of grade G_i will be within the specified tolerance. A more reduced quality descriptor for the gray scale can be defined as the average of $K(G_i)$ across all grades:

$$\bar{K} = \frac{1}{8} \sum_{i=1}^8 K(G_i) \quad (2)$$

where the summation goes up to 8 (not 9) because color pair 5 was excluded (two identical physical samples).

The calculated values of \bar{K} are shown in Fig. 5 (AATCC) and Fig. 6 (SDCE), where each subplot corresponds to a given incidence angle, θ_i , and the values are plotted for different reflection angles, θ_s .

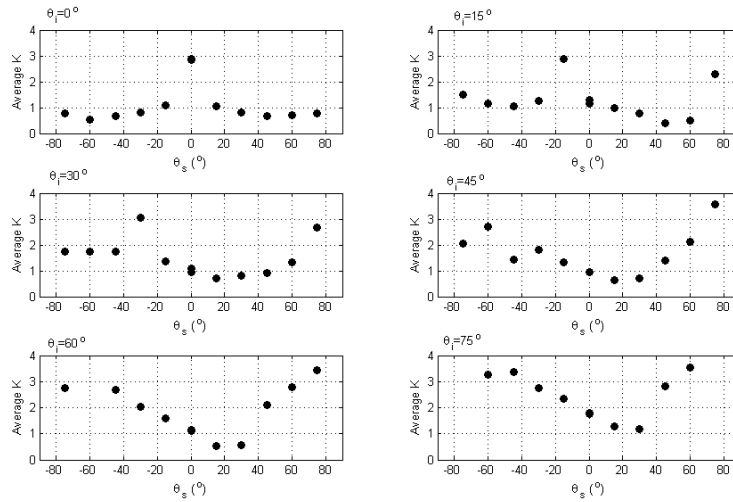


Fig. 5: Average K value [see Eqs. (1) and (2)] for every measurement geometry of the AATCC gray scale for color change.

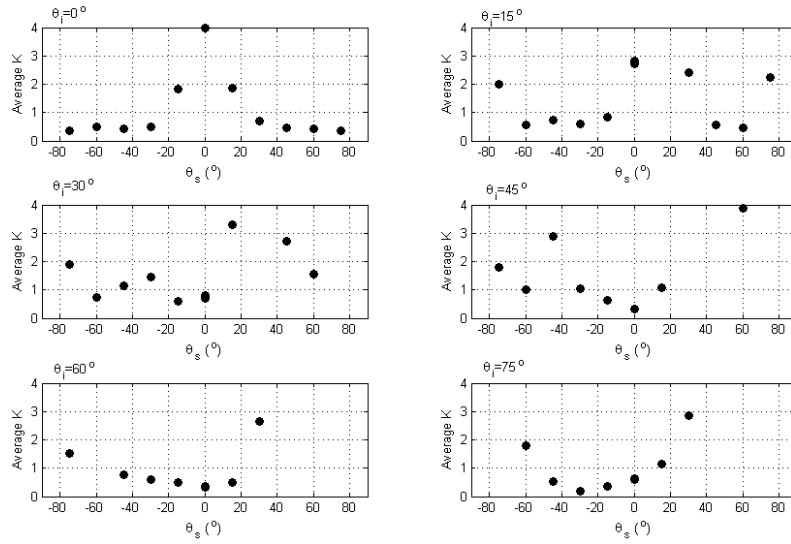


Fig. 6: Same as in Fig. 5, but for the SDCE gray scale for color change.

As expected, the results in Fig. 5 and Fig. 6 suggest that the AATCC and SDCE gray scales are acceptable (i.e. $\bar{K} < 1$) at geometry $45^\circ:0^\circ$. However, they worsen the farther they are from such geometry. In general, it holds that $\bar{K} < 1$ for the SDCE gray scale at low incidence angles ($\theta_i = 0^\circ$ and 15°) and most viewing angles, but it doesn't at large reflection angles in combination with moderate and large incidence angles. In addition, the performance is worse at low aspecular angles. This result is consistent with ASTM D2616-12 recommendation [see 8.1.2 in ref. 5] indicating to “avoid specular reflection of the source by illuminating the specimens at an angle of about 45° and viewing them perpendicularly, or the reverse geometry”. However, \bar{K} does not depend so much on the aspecular angle in the case of the AATCC gray scale. Whereas the performance of the AATCC gray scale is less dependent on the geometry, the SDCE gray scale provides a better match with the standard values at the standard geometry $45^\circ:0^\circ$.

For the studied gray scales, the value of \bar{K} , lower or slightly larger than the unity, suggests that it is acceptable to use them within a difference of around 15° both for incidence and viewing angles at $45^\circ:0^\circ$, but not for much larger differences. In agreement with the ASTM recommendation mentioned in the previous paragraph, there are geometries that should be definitively avoided when using these gray scales. They are those with an abrupt variation of the BRDF with respect to values at geometries $0^\circ:45^\circ$ and $45^\circ:0^\circ$, corresponding to large incidence and viewing angles (see Fig. 2 and Fig. 3).

The conclusion that these scales should not be used with angular differences larger than 15° with respect to the incidence and viewing standard directions must be considered in visual experiments. For instance, it involves that, given the distance between the centers of the more extreme pairs in the scales (17-18 cm), they should be observed from a minimum distance of around 34 cm; or that the angle subtended by the light source irradiating the scale should not be larger than 30° , which must be considered when using these scales in color-assessment cabinets.

Results showed that, at large incidence and viewing angles, not only the color differences are very different to the standard values (see Fig. 5 and Fig. 6, where some values are out of scale), but also some lightness “inversions” are frequently found. That is, at these extreme geometries some chips have lightness differences with different signs than those reasonably expected from their nominal grades. Inversions are more usual for dark pairs (low grade values), and can be identified from Fig. 2 and Fig. 3 as cases where the BRDF curves corresponding to two different chips intersect.

The results in Fig. 5 and Fig. 6 are average values of $K(G_i)$ for the color pairs in the gray scales, not values for individual color pairs. Assuming CIE D65 standard illuminant and CIE 1964 standard colorimetric observer, CIELAB color differences for each of the color pairs in both gray scales at all measured geometries can be found at the *Appendix*. CIELAB color differences are also plotted in Fig. 7 for the color pairs in the AATCC and SDCE gray scales at the 6 geometries implemented in some commercial multi-angle spectrophotometers and color-assessment cabinets (e.g. BYK-mac and byko-spectra effect cabinet). For reference, the standardized values for each color pair at geometry $45^\circ:0^\circ$ with their corresponding tolerances (Table 1) are shown as error bars in Fig. 7. The results for geometries $45^\circ:20^\circ$ and $45^\circ:-65^\circ$ were not experimentally measured, but interpolated from spectral radiance factors at nearest measured geometries. From Fig. 7, all color pairs in the SDCE gray scale are within the standard tolerances fixed for the geometry $45^\circ:0^\circ$, whereas pairs 3-4 and 4-5 for the AATCC gray scale are slightly outside of such tolerances [3-5]. In addition, it can be observed that for geometries with low aspecular angles ($45^\circ:60^\circ$, $45^\circ:30^\circ$, and even $45^\circ:20^\circ$) the discrepancies with respect to the standard values (Table 1) are considerably higher in the SDCE than in the AATCC gray scale, which is in agreement with previous results shown in Fig. 5 and Fig. 6.

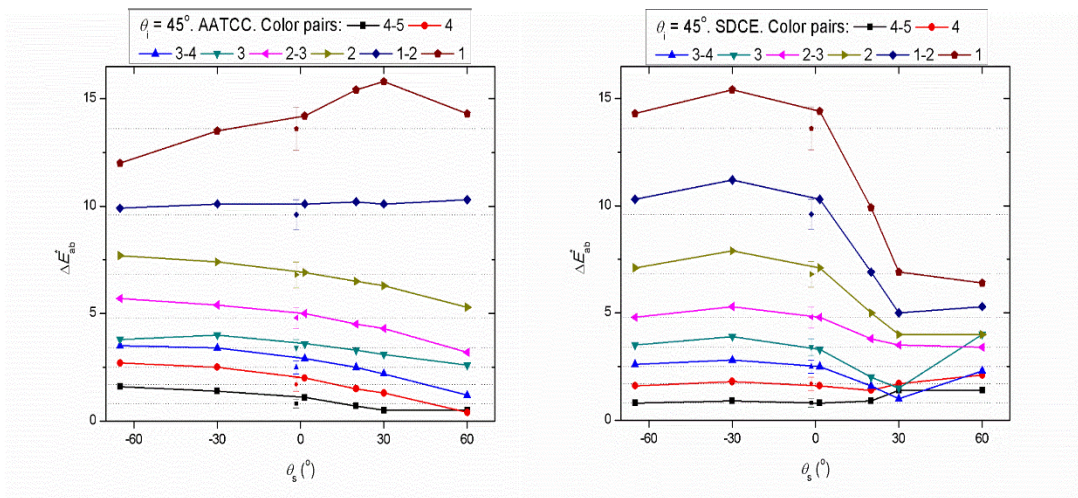


Fig. 7. CIELAB color differences (illuminant D65 / CIE 1964 observer) for the color pairs in the AATCC (left) and SDCE (right) gray scales, for the geometries implemented in some commercial multi-angle spectrophotometers and color-assessment cabinets. The results for $\theta_s = 0^\circ$ were slightly displaced to the right, to plot error bars indicating standard tolerances (Table 1, [3-5]).

4. Conclusions

Spectral BRDF was measured for gray scales manufactured by the AATCC and SDCE, which are typically used in color-fastness tests in the textile industry and other visual applications. For the sake of representativeness, two recently-acquired specimens of each gray scale were measured. It has been shown that both, the lightness of individual chips and the color

difference of pairs in these gray scales depend on the irradiation and detection directions, and the magnitudes of these goniochromatic effects have been quantified. Inversion (the lighter becomes the darker) or negligible lightness difference in some pairs have been observed at some geometries. According to the measurements, the chips in the AATCC gray scale present a lower dependence on illumination and viewing directions than those in the SDCE gray scale. The geometries at which some color pairs of these gray scales are not within the tolerances standardized by AATCC, ISO, and ASTM (which recommend a $0^\circ:45^\circ$ or $45^\circ:0^\circ$ geometry) have been evaluated, and correspond to abrupt variations of the BRDF with respect to the aforementioned geometries. It has been concluded that, whereas the color differences of the AATCC gray scale are less dependent on the geometry, the SDCE gray scale provides a slightly better match with the standard values at the standard geometry $45^\circ:0^\circ$. For tests or visual assessments illuminating a gray scale with an extended source, for which many irradiation directions are involved, it is advisable to perform direct spectroradiometric measurements of the gray chips as viewed by real observers, avoiding the assumption of the standardized values of color differences for the $0^\circ:45^\circ$ or $45^\circ:0^\circ$ geometries (i.e. Table 1). The data shown in this paper indicate that the goniochromatic properties of the AATCC and SCDE gray scales for color change are considerably different, and should help to prevent bad practices in future visual assessments involving gray scales. These scales should not be used with angular differences larger than 15° with respect to the incidence and viewing standard directions ($45^\circ:0^\circ$), and this must be considered in visual experiments.

5. Funding and acknowledgments

5.1 Funding

Jointly funded by the EMRP participating countries within EURAMET and the European Union (EMRP IND52); Ministerio de Economía y Competitividad (FIS2013-40661-P, FIS2016-80983-P); Comunidad de Madrid (S2013/MIT-2790).

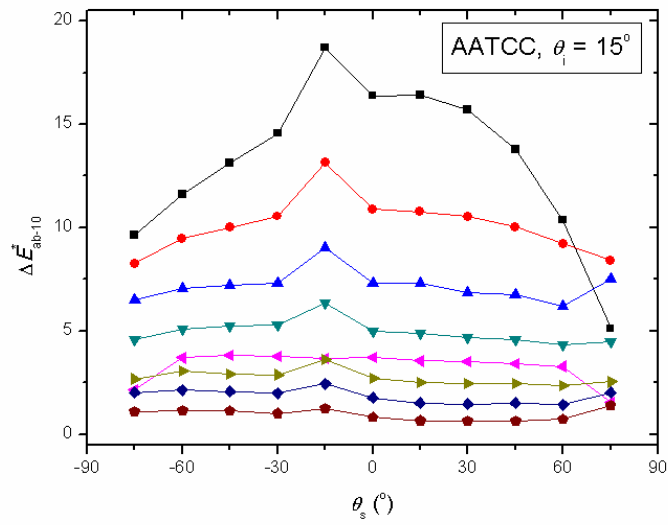
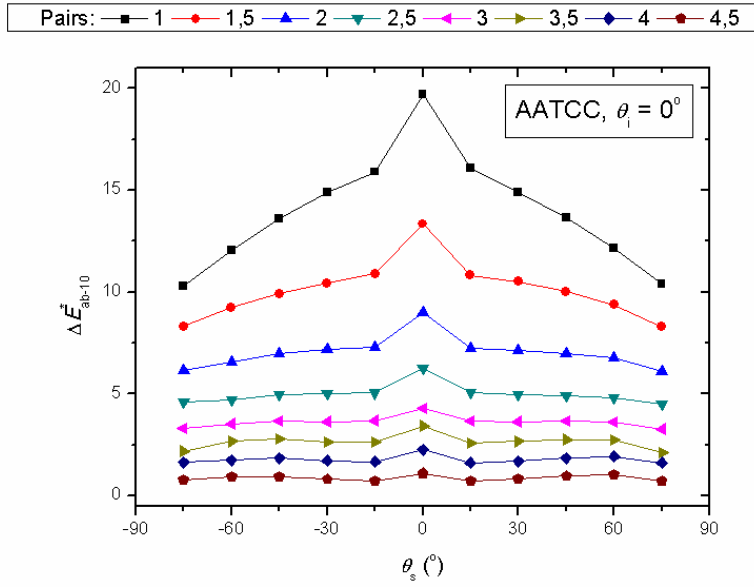
5.2 Acknowledgments

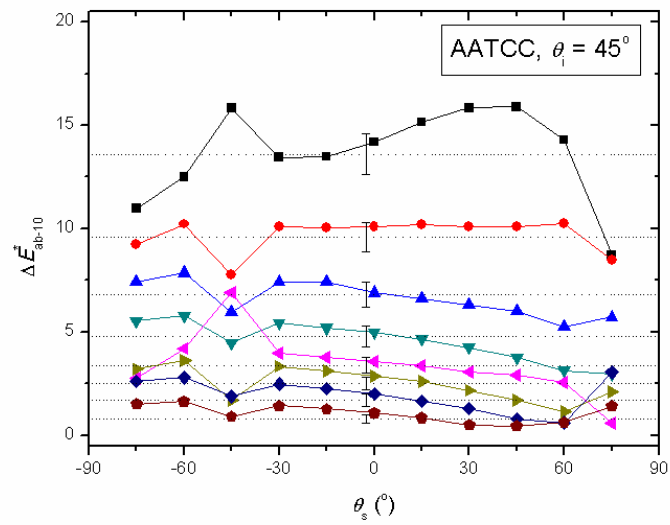
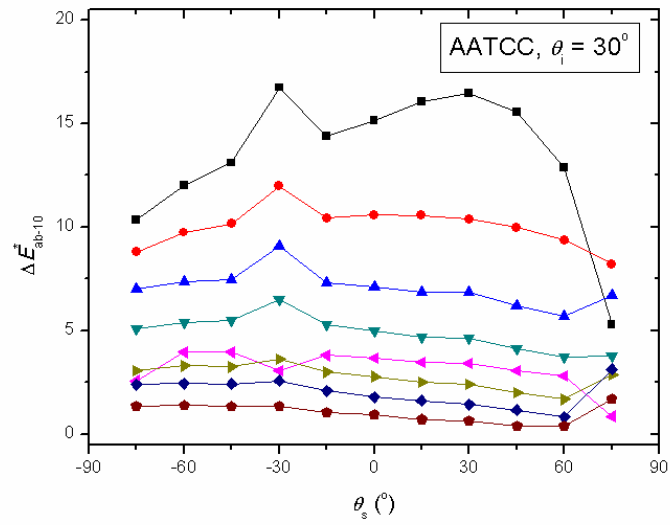
This work was supported by the Ministry of Economy and Competitiveness of the National Government of Spain, Research Project FIS2016-80983-P, with support from European Regional Development Fund (ERDF), by Community of Madrid (project SINFOTON-CM: S2013/MIT-2790), and by the EMRP IND52 Project “Multidimensional reflectometry for industry” (xD-Reflect). The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.

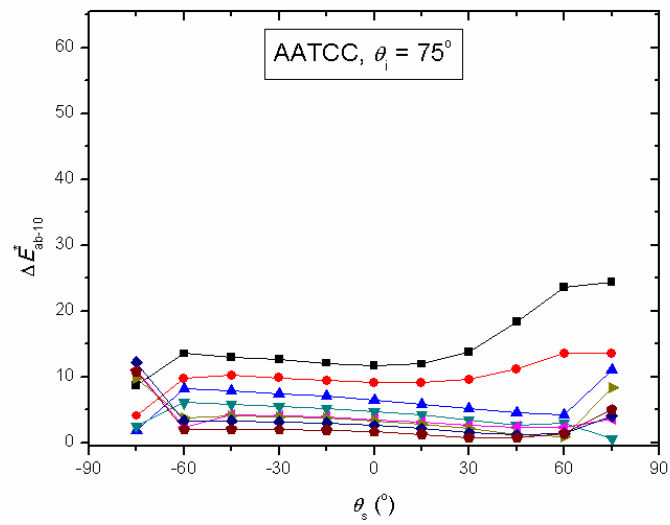
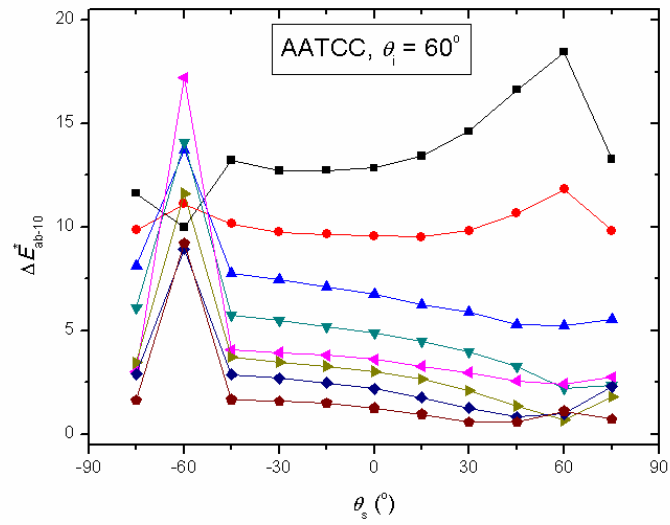
5.3 Disclosures

The authors declare no conflicts of interest.

Appendix: Color differences of the pairs in the studied gray scales.

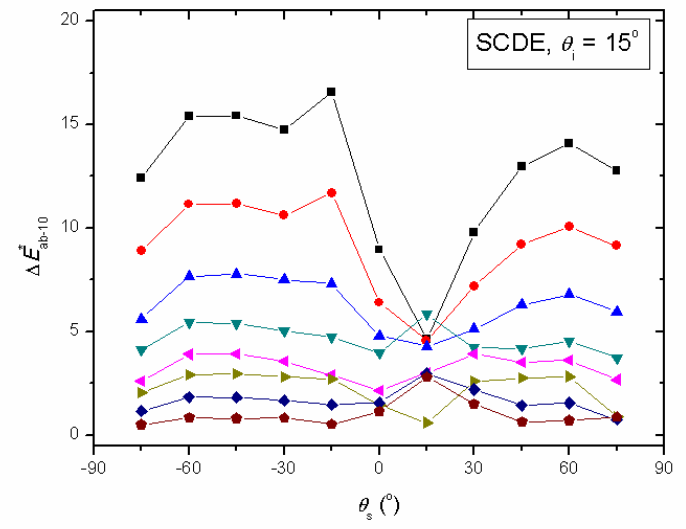
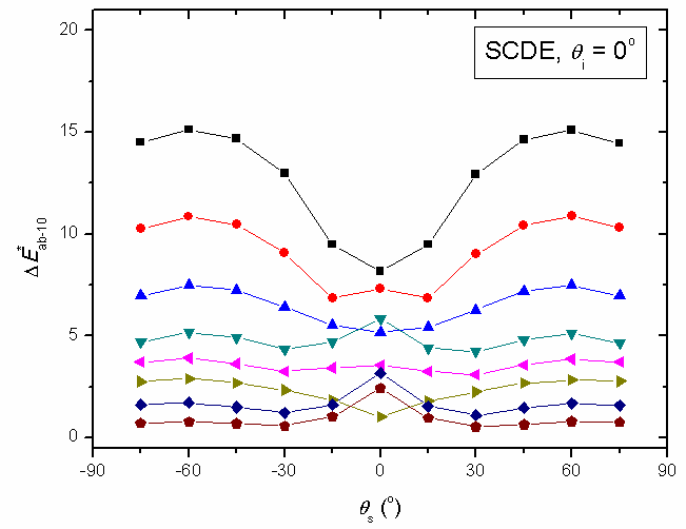


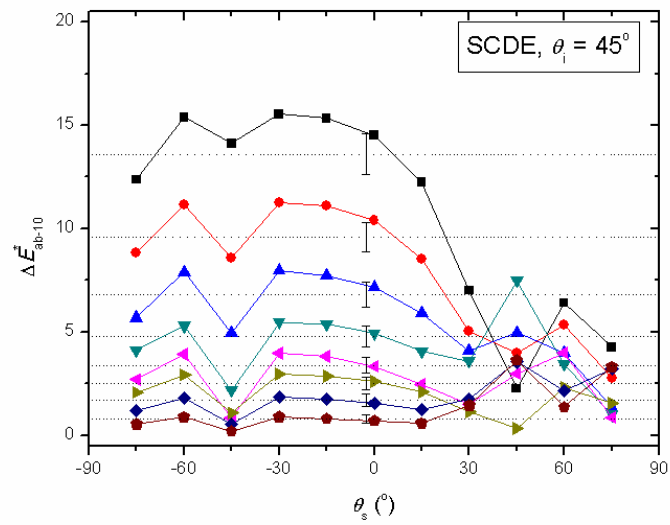
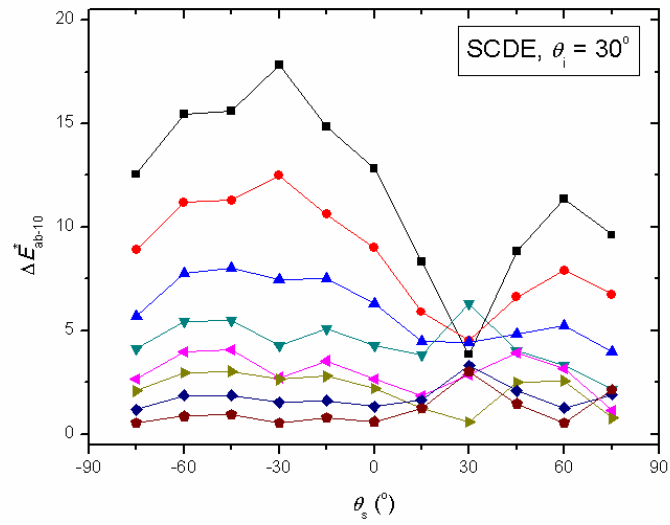


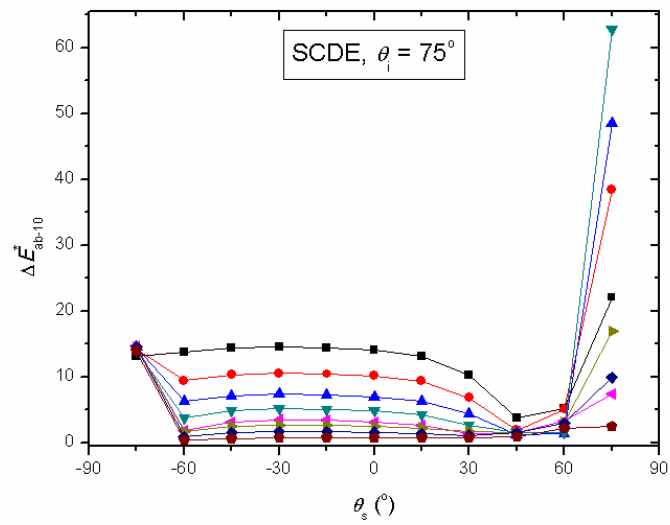
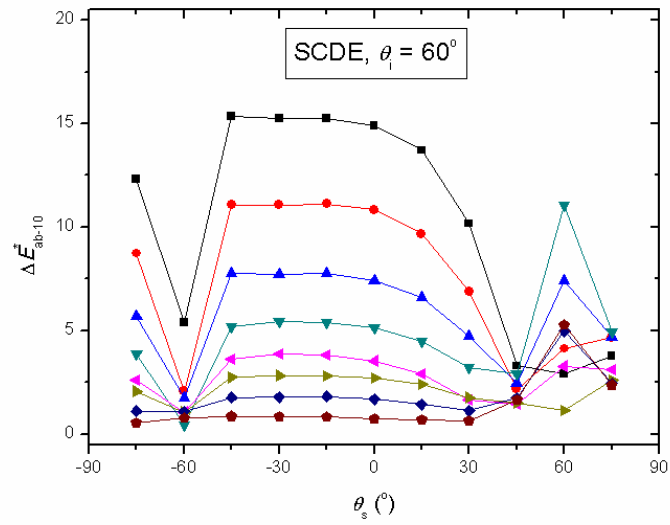


AATCC gray scale for color change. Color differences of the pairs at different geometries, calculated with CIE D65 standard illuminant and CIE 1964 standard colorimetric observer.

Pairs: —■— 1 —●— 1,5 —▲— 2 —▼— 2,5 —◆— 3 —▶— 3,5 —◆— 4 —●— 4,5







SDCE gray scale for color change. Color differences of the pairs at different geometries, calculated with CIE D65 standard illuminant and CIE 1964 standard colorimetric observer.

References

1. G. Klein, *Industrial Color Physics*. New York: Springer, 2010, p. 192.
2. Fastness Tests Co-ordinating Committee (F.T.C.C.), *Publication XI. The development of the geometric grey scales for fastness assessment*, *J. Soc. Dyers Colourists* **69**, 404–409 (1953).
3. AATCC Evaluation Procedure 1-2012. Gray Scale for Color Change. AATCC Technical Manual, 2015, pp. 415-416.
4. ISO 105-A02, Textiles-Test for colour fastness- Part A02: Grey scale for assessing change in colour.
5. ASTM D 2616 – 12, Standard Test Method for Evaluation of Visual Color Difference With a Gray Scale.
6. M. R. Luo, and B. Rigg, “BFD(l:c) colour-difference formula. Part I – Development of the formula,” *J. Soc. Dyers Colour.* **103**, 86-94 (1987).
7. S.-S. Guan and M. R. Luo, “Investigation of parametric effects using small colour differences,” *Color Res. Appl.* **24**, 331-343 (1999).
8. E. D. Montag and D. C. Wilber, “A comparison of constant stimuli and gray-scale methods of color difference scaling,” *Color Res. Appl.* **28**, 36-44 (2003).
9. S. G. Kandi and M. A. Tehran, “Investigating the effect of texture on the performance of color difference formulae,” *Color Res. Appl.* **35**, 94-100 (2010).
10. R. Shamey, R. Cao, T. Tomasino, S. S. H. Zaidy, K. Iqbal, J. Lin and S. G. Lee, “Performance of select color-difference formulas in the blue region,” *J. Opt. Soc. Am. A* **31**, 1328-1336 (2014).
11. M. Melgosa, J. Martínez-García, L. Gómez-Robledo, E. Perales, F. M. Martínez-Verdú and T. Dauser, “Measuring color differences in automotive samples with lightness flop: A test of the AUDI2000 color-difference formula,” *Opt. Express* **22**, 3458-3467 (2014).
12. O. Gómez, E. Perales, E. Chorro, F. J. Burgos, V. Viqueira, M. Vilaseca, F. M. Martínez-Verdú and J. Pujol, “Visual and instrumental assessments of color differences in automotive coatings,” *Color Res. Appl.* **41**, 384-391 (2016).
13. A. Mohammadalizadeh, F. Ameri, S. Moradian, M. Melgosa, F. Mirjalili. Modelling the orange-peel texture for chromatic and achromatic samples. *Prog. Org. Coat.* **135**, 148-155 (2019).
14. CIE 015:2018, Colorimetry, 4th Edition. Technical Report. CIE Central Bureau, Vienna, 2004.
15. M. Melgosa, L. Gómez-Robledo, G. Cui, C. Li, A. Ferrero, B. Bernad, J. Campos, N. Richard, and C. Fernández-Maloigne, “Using a standard grey scale for colour change in a multi-angle colour assessment cabinet,” in *Proceedings of the 4th CIE Expert Symposium on Colour and Visual Appearance*, ed. (CIE, 2016), pp. 287-294.
16. B. Bernad, A. Ferrero, C. Strothkamper, C. J. Campos, A. Pons, T. Quast, T. K.O. Hauer, A. Schirmacher. Deviation of white diffuse reflectance standards from perfect reflecting diffuser a visible and near-infrared spectral ranges. *Metrologia* **56**, 055005 (2019).
17. A. M. Rabal, A. Ferrero, J. Campos, J. L. Fontecha, A. Pons, A. M. Rubiño, and A. Corróns, “Automatic gonio-spectrophotometer for the absolute measurement of the spectral BRDF in in- and out-of-plane and retroreflection geometries,” *Metrologia* **49**, 213–223 (2012).
18. Joint ISO/CIE Standard, ISO/CIE 11664-6:2014(E), Colorimetry – Part 6: CIEDE2000 Colour-Difference Formula.
19. M. Melgosa, E. Hita, J. Romero, L., and Jiménez del Barco, “Some classical color differences calculated with new formulas,” *J. Opt. Soc. Am. A* **9**, 1247-1253 (1992).
20. M. Huang, H. Liu, G. Cui, M. R. Luo, and M. Melgosa, “Evaluation of threshold color differences using printed samples,” *J. Opt. Soc. Am. A* **29**, 883-891 (2012).
21. L. M. Cárdenas, R. Shamey, and D. Hinks, D., “Development of a novel linear gray scale for visual assessment of small color differences,” *AATCC Review* **9**, 42-47 (2009).
22. N. Dekker, M. Lucassen, E.J.J. Kirchner, P. Urban, and R. Huertas, “Mathematical limitations when choosing psychophysical methods: geometric versus linear grey scales”, in: M.V. Ortiz Segovia, P. Urban, J.P. Allebach (Eds.), *Proc. SPIE 9018, Conference on Measuring, Modeling, and Reproducing Material Appearance (MMRMA, San Francisco, 2 February 2014)*.
23. F. E. Nicodemus, J. C Richmond, J. J Hsia, I. W. Ginsberg, and T. Limperis, “Geometrical considerations and nomenclature for reflectance,” *NBS Monograph 160*, National Bureau of Standards, Gaithersburg, MD, 1997.

24. Accurate Measurement of Optical Properties of Materials Edited by Thomas A. Germer, Joanne C. Zwinkels, and Benjamin K. Tsai, 2014 Elsevier Inc.
25. A. Ferrero, A. M. Rabal, J. Campos, A. Pons, and M. L. Hernanz, "Spectral and geometrical variation of the bidirectional reflectance distribution function of diffuse reflectance standards," *Appl. Opt.* **51**, 8535-8540 (2012).
26. R. S. Hunter, and R. W. Harold, *The measurement of appearance*, Reston: John Wiley & Sons, 1987, p. 78.
27. A. Ferrero, B. Bernad, J. Campos, N. Richard, C. Fernández-Maloigne, and M. Melgosa, "Spectrally-averaged BRDF measurements corresponding to the 9 chips of the gray scales AATCC and SDCE at six different irradiation angles," *figshare* (2020), <https://osapublishing.figshare.com/s/7ccd11e7388133f13a46>.