

Repeatability and reproducibility of specular gloss meters in theory and practice

Frédéric Bernard Leloup, Jan Audenaert, Gaël Obein, Guillaume Ged, Peter Hanselaer

© American Coatings Association 2016

Abstract In order to numerically quantify and to ensure consistency of the visual appearance of the end product, many industries perform color and gloss measurements, through the results of which they define maximum tolerances during quality control. For this, they may primarily rely on precision data for repeatability and reproducibility as reported by instrument manufacturers, who develop their instruments according to standardized measurement geometries defined in international standards. This paper reports on an investigation of the inter-instrument agreement between specular gloss meters which conform to universally adopted gloss standards. Six commercially available instruments, manufactured by three different companies, and twenty-five gloss artifacts, with specular gloss values ranging between 2 and 110 gloss units (60° geometry), were selected for use in the study. The repeatability and reproducibility of the instruments were assessed according to the criteria described in ASTM D523-14 and ISO 2813:2014, and to the specifications reported by the instrument manufacturers. Ray tracing simulations were performed by use of

the bidirectional reflectance distribution function measurement data of two samples, and based on the standardized optical design of the specular gloss meter, in order to get a better insight on the expected tolerances of the gloss readings from a theoretical point of view. On average, the practical results indicate that both the repeatability and reproducibility values are higher than those specified in the recommendations and by the manufacturers. Ray tracing simulations confirm this finding. While specified repeatability and reproducibility thresholds are based on ideal standards, in practice, there are parameters that can lead to significant higher deviations. Care should therefore be taken when analyzing measurement results obtained from different instruments, and when defining tolerances for evaluation of gloss measurements.

Keywords Specular gloss, Gloss meter, Inter-instrument agreement

Introduction

Surface gloss, together with surface color and surface texture, is of particular importance in several industrial applications for which the final appearance determines the quality impression of the end product (e.g., automobile, paint, steel, and packaging).¹ Hence, besides a visual inspection, quality control of surface gloss is generally performed using a so-called specular gloss meter. This commercial optical device quantifies gloss from the degree of specular reflection from the surface, in comparison to the degree of specular reflection from a black glass standard, as measured in a particular measurement geometry.²

In order to make a comparison between results obtained from different instruments, the measurement geometries as well as the entire optical layout of the specular gloss meter has been standardized. General

F. B. Leloup (✉), J. Audenaert, P. Hanselaer
Light & Lighting Laboratory, Department of Electrical Engineering (ESAT), KU Leuven, Gebroeders De Smetstraat 1, 9000 Ghent, Belgium
e-mail: Frederic.Leloup@kuleuven.be

J. Audenaert
e-mail: Jan.Audenaert@kuleuven.be

P. Hanselaer
e-mail: Peter.Hanselaer@kuleuven.be

G. Obein, G. Ged
Conservatoire National des Arts et Métiers, 292 Rue Saint-Martin, 75141 Paris Cedex 03292, France
e-mail: Gael.Obein@lecnam.net

G. Ged
e-mail: Guillaume.Ged@lecnam.net

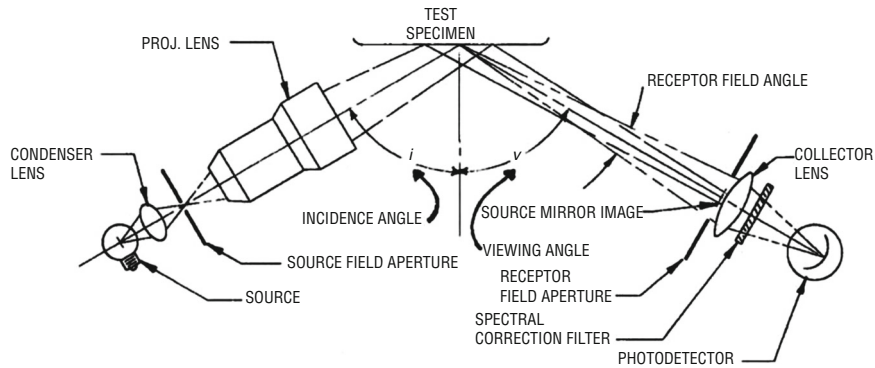


Fig. 1: Optical design of a converging-beam specular gloss meter, according to the specifications defined in ASTM D523-14³

Table 1: Instrument specifications of the six specular gloss meters used in the study

Manufacturer	BYK-Gardner		Rhopoint		Zehntner	
Instrument type	Micro-TRI-gloss	Micro-TRI-gloss-S	IQ Flex 20	IQ Trigloss	ZGM 1110	ZGM1120
Measurement geometry	20°/60°/85°	20°/60°/85°	20°	20°/60°/85°	20°/60°/85°	20°/60°/85°
Repeatability (GU)	±0.2	±0.2 (±0.1) ^a	±0.2	±0.2	±0.1	±0.1
Reproducibility (GU)	±0.5	±0.5 (±0.2) ^a	±0.5	±0.5	±0.5	±0.5
Traceability	BAM	BAM	BAM	BAM	BAM	BAM

^a Repeatability and reproducibility in the range of 0–10 GU (60° geometry)

test methods for the measurement of the specular gloss of nonmetallic surfaces, i.e., ASTM D523-14 and ISO 2813:2014,^{3,4} specify three measurement geometries (20°, 60°, and 85° geometry) according to the angle of incidence and the angle of view at which the reflected flux is recorded. As an example, the optical layout of the converging-beam measurement geometry, as defined in ASTM D523-14, is depicted in Fig. 1.

Accurate control of the effective directions of illumination and viewing, and of the field angles of the incidence beam and viewing beam, is accomplished by means of stops, lenses, and mirrors, of which the dimensions and tolerances are also defined.

Instrument manufacturers commonly state that their instruments comply with the standard methods for the measurement of specular gloss. Yet, apart from some older studies,^{5–7} little has been reported about the agreement between gloss measurement results obtained with gloss meters from different instrument manufacturers over an extended glossiness range. Sève demonstrated from a theoretical point of view that the ratio of specular gloss values of a high-gloss and a low-gloss surface is not only determined by the two materials, but also by the solid angles used in the measuring apparatus.⁶ Even for instruments that conform to the same standard and geometry, different gloss readings may be obtained due to the different relative dimensions of the source and receptor apertures. Arney et al. provided a practical example, by comparing gloss values measured with a BYK-Gardner

gloss meter and an Ihara gloss meter, on a variety of paper samples, printed images, and plastic samples.⁷ Both instruments adopt a 75° geometry, and conform to the TAPPI standard T480 for the measurement of paper gloss.⁸ Yet, for each sample, the BYK-Gardner readings were higher than the Ihara readings. Moreover, no linear correlation was found between the results obtained with both apparatus. In agreement with Sève, Arney et al. conjectured this phenomenon to result from the difference in size of the receptor angle, which appeared to be 2.9° for the BYK-Gardner instrument and only 2° for the Ihara instrument.

In this paper, the repeatability and reproducibility of specular gloss meters are investigated from a theoretical as well as a practical point of view. Specular gloss measurements are performed using six different commercial gloss meters and 25 test samples. The repeatability and reproducibility of the test instruments are evaluated through the average values of the readings, according to the criteria described in ASTM D523-14 and ISO 2813:2014, and to the specifications reported by the instrument manufacturers. The correlation between the measurement results obtained with the different gloss meters is assessed. To better understand the tolerances and deviations of the experimental gloss measurement results, ray tracing simulations are performed by use of the bidirectional reflectance distribution function data of two test samples, and based on the specifications of the optical design of the gloss meter as described in the related standards.



Fig. 2: Example picture of a white glossy (left) and a white semi-gloss (right) sample, developed as part of a gloss scale containing 40 artifacts. Both samples are illuminated with a uniform light source positioned above the samples. The left sample obviously possesses a higher reflectance; the lightness of the sample is higher and the image of the character in the sample surface can be better discerned

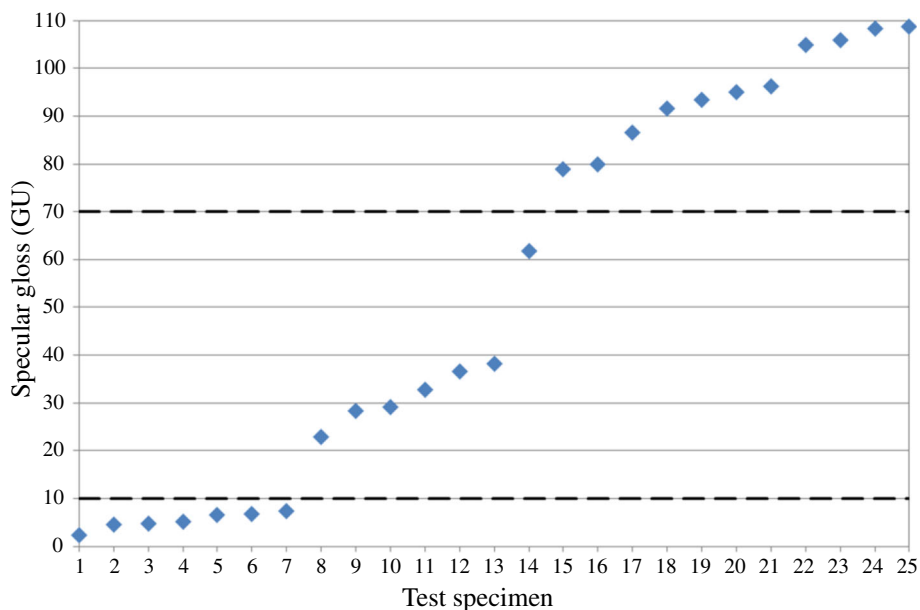


Fig. 3: Average gloss values of the 25 test samples, expressed in GU, as obtained from all instrument readings (except Rhopoint IQ Flex 20) in the 60° measurement geometry

Practical approach

Devices under test

Suppliers of commercial specular gloss meters were contacted with the request to provide calibrated devices for testing. Six types of specular gloss meters from three different instrument manufacturers were received. A description of the main specifications of each device, as provided by the manufacturers, is given in Table 1. With the exception of one instrument, which only measures in the 20° geometry (Rhopoint IQ Flex 20), each device measures the specular gloss in three specular gloss measurement geometries, i.e., the 20°, 60°, and 85°

geometry. Both the stated repeatability and reproducibility values, derived by use of ideal standards according to the definitions described in the normalization standards (see further), are expressed in gloss units (GU). Each instrument is calibrated against a reference standard traceable to the BAM Federal Institute for Materials and Testing (Germany).

Test samples

The test samples selected for use in this study originate from a newly developed gloss scale as part of a joint research project entitled “Multidimensional Reflec-

Table 2: Maximum acceptable gloss differences (expressed in GU), for evaluation of the repeatability and reproducibility, respectively, of painted samples according to ASTM D 523-14 and ISO 2813:2014

Measurement geometry	Repeatability threshold value (GU)	Reproducibility threshold value (GU)	
		ASTM	ISO
20°	1.7	6.4	4
60°	0.9	3.5	3
85°	0.8	7.2	2

For ISO 2813:2014 only the reproducibility threshold values are reported, since the ISO method for evaluating the repeatability is based on the absolute difference between mean values of 2 separate sets of 3 values, and not on the difference between single determinations which was used in this study

Table 3: Average repeatability values for each sample, expressed in GU

Sample	Repeatability (GU)	Sample	Repeatability (GU)	Sample	Repeatability (GU)
1	0.7	10	12.6	19	1.0
2	0.4	11	17.1	20	4.1
3	3.0	12	3.1	21	3.6
4	3.3	13	4.4	22	2.2
5	4.1	14	7.0	23	3.2
6	2.7	15	8.3	24	3.8
7	6.4	16	5.7	25	4.4
8	11.3	17	0.9		
9	11.2	18	2.4		

Values exceeding the threshold values reported in Table 2 are indicated in bold

tometry for Industry” (xDRreflect), performed by a consortium of eight national metrology institutes, two universities, and one industrial partner, within the framework of the European Metrology Research Programme (EMRP) of the European Association of National Metrology Institutes (EURAMET). The sample set contains 40 items (product dimensions 9.5 × 9.5 cm), all made of glass, but differing according to one or more of the following three parameters: lightness (white, gray or black), refractive index (1.47 vs 1.53), and roughness (three levels). From the entire sample collection, 25 artifacts with specular gloss values ranging between 2 and 110 GU in the 60° measurement geometry, were selected based on a visual inspection of the surface uniformity. By way of example, a picture showing the specular reflection from two white samples (glossy vs semi-glossy) is presented in Fig. 2.

Practical results

Measurement protocol

For each combination of specular gloss meter and test sample, a measurement on five representative positions was performed in the three measurement geometries (except for the Rhopoint IQ Flex 20 for which only the 20° measurement geometry was taken into account). The uniformity and directionality of the samples were

assessed by taking two out of the five measurements with the gloss meters rotated over 90°.

Repeatability and reproducibility were checked in the recommended measurement geometry according to the average gloss values obtained from all instruments in the 60° geometry. The 20° geometry was adopted for samples with an average gloss value in the 60° geometry being larger than 70 GU, the 60° geometry was chosen for samples with an average gloss value in the 60° geometry lying between 10 and 70 GU, and the 85° geometry was selected for the samples with an average gloss value in the 60° geometry being lower than 10 GU.

Repeatability is defined as the agreement between two results from single determinations obtained on the same sample with the same instrument.³ Therefore, for each gloss meter, the repeatability was investigated from the difference between the maximum and minimum of the five recorded readings on each sample.

Reproducibility is defined as the agreement between two results, each being the average of the five consecutive readings* with one instrument, obtained on the same sample by use of different instrumentation.³ As such, reproducibility was assessed both from the difference between the maximum and minimum average gloss value obtained with all six gloss meters on

* ASTM recommends assessing the reproducibility from two results which are the mean of 3 readings instead of 5.

Table 4: Average reproducibility values for each sample, expressed in GU

Sample	Reproducibility (GU)	Sample	Reproducibility (GU)	Sample	Reproducibility (GU)
1	0.8	10	<u>10.7</u>	19	<u>13.9</u>
2	0.9	11	<u>4.1</u>	20	<u>14.3</u>
3	1.9	12	1.6	21	<u>9.2</u>
4	1.4	13	<u>6.2</u>	22	<u>7.7</u>
5	1.8	14	<u>6.1</u>	23	<u>7.2</u>
6	1.8	15	3.2	24	<u>8.5</u>
7	2.7	16	<u>8.9</u>	25	<u>9.1</u>
8	2.2	17	<u>9.9</u>		
9	3.3	18	<u>10.5</u>		

Values only exceeding the ISO threshold values reported in Table 2 are indicated in bold. Values exceeding both the ASTM and ISO threshold values reported in Table 2 are indicated in bold and underlined

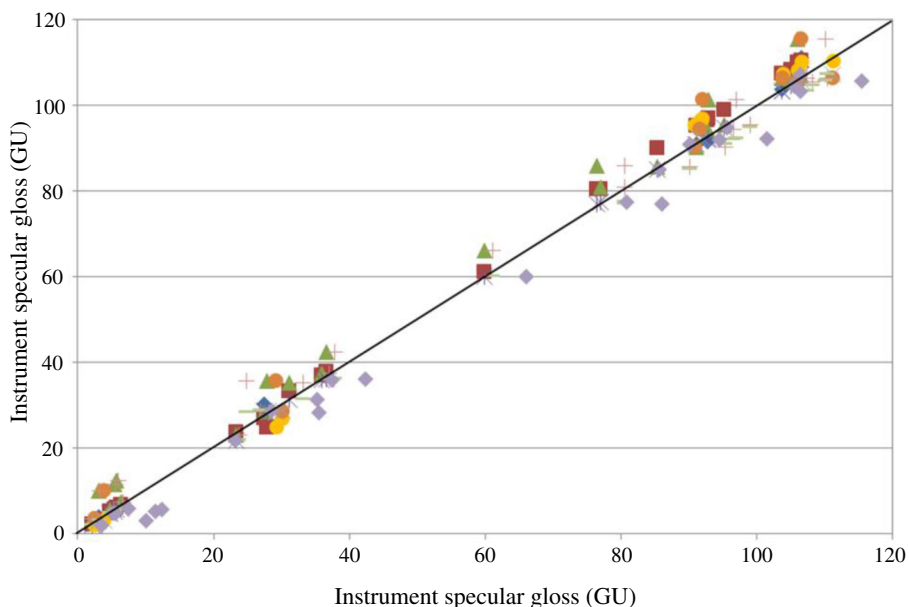


Fig. 4: Average specular gloss of each sample for each gloss meter tested, presented against the average specular gloss as obtained with another gloss meter, both expressed in GU. Each combination of symbol and color represents a different pair of gloss meters

each sample, and from the difference between average gloss values for each combination of two gloss meters.

Both the recorded repeatability and reproducibility values were compared to the threshold values as described in ASTM D523-14 and ISO 2813:20014 (see further), and to the specifications reported by the instrument manufacturers (see Table 1).

Measurement results

The average gloss values of the 25 test samples, as obtained from all instrument readings (except Rho-point IQ Flex 20) in the 60° measurement geometry, are illustrated in Fig. 3. The gloss values range from 2 to 110 GU. Seven out of the 25 samples’ average gloss

value is below 10 GU, while 11 samples’ average value is larger than 70 GU. For these samples, the gloss readings obtained in the 85° and 20° geometry, respectively, were used further in the study, in agreement with the recommendations made in ASTM D523-14 and ISO 2813:2014.

With the exception of four samples, the average coefficient of variation (i.e., the relative standard deviation of the five consecutive measurements performed with each instrument on each sample) is below a few percent, indicating a good uniformity of the gloss artifacts, free from any directionality. For samples 8, 9, 10, and 11, however, the coefficient of variation numbers 20%, 15%, 21%, and 21%, respectively. Obviously, the development of uniform samples in the low/mid-gloss range (between 20 and 30 GU—60°

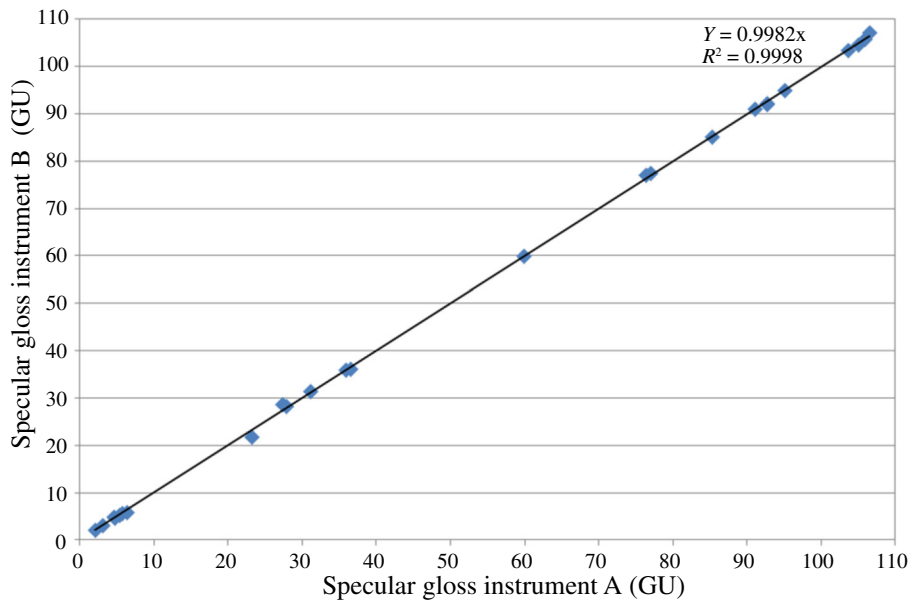


Fig. 5: Average specular gloss of each sample measured with instrument A, presented against the average specular gloss measured with instrument B, both expressed in GU. Instrument A and B are the two instruments for which the data agree best; both instruments are developed by the same manufacturer

Table 5: Reproducibility for each sample, expressed in GU, as obtained by comparison of instrument C and instrument D

Sample	Reproducibility (GU)		Sample	Reproducibility (GU)		Sample	Reproducibility (GU)	
	Before	After		Before	After		Before	After
1	0.0	0.3	10	9.9	3.2	19	0.7	2.1
2	0.3	0.0	11	8.7	3.3	20	1.7	1.2
3	0.4	1.4	12	8.9	3.0	21	10.7	4.8
4	0.0	0.4	13	10.3	2.5	22	2.1	1.2
5	1.2	0.7	14	9.1	6.8	23	0.5	1.5
6	1.3	0.6	15	7.7	7.1	24	4.5	2.3
7	2.6	0.6	16	5.5	9.0	25	5.0	0.7
8	2.3	5.1	17	8.3	8.5	Average	4.5	3.2
9	2.1	7.2	18	9.1	7.1			

Values only exceeding the ISO threshold values reported in Table 2 are indicated in bold. Values exceeding both the ASTM and ISO threshold values reported in Table 2 are indicated in bold and underlined

geometry) of the gloss scale seems most complicated. The resulting data of these four samples are nevertheless further analyzed, keeping in mind that the lack of uniformity influences the repeatability and reproducibility results obtained for these samples.

On the basis of studies performed by several laboratories, ASTM and ISO defined threshold values for judging the acceptability of results at the 95% confidence level. According to ASTM, the repeatability should be considered suspect if the difference between two readings is larger than the values stated in Table 2. These values are higher than the threshold values reported by the instrument manufacturers (cf. Table 1). Average repeatability values, i.e., the average of the calculated repeatability of all gloss meters,

are presented for each sample in Table 3. As expected, average repeatability values of samples 8–11 are extremely high due to the lack of uniformity of these samples, and caution should be exercised when evaluating these numbers. However, even neglecting these data, with the exception of the values of 4 samples (sample 1, 2, 17, and 19), all average repeatability values exceed the repeatability thresholds indicated in ASTM D 523-14, compared in the appropriate measurement geometry. For each individual gloss meter, the repeatability thresholds are exceeded for at least 20 out of the 25 test samples, while the manufacturers' repeatability threshold values are never reached.

Similar to the repeatability, ASTM and ISO defined threshold values for judging the acceptability of results

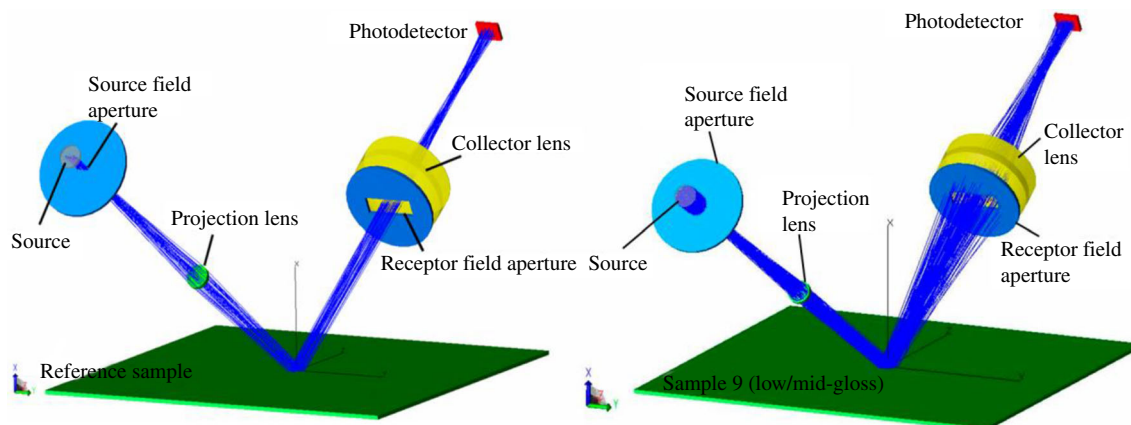


Fig. 6: Example image of a performed ray tracing simulation on the black glass reference sample (left), and on the low/mid-gloss sample (right), depicting the optical components of the converging-beam gloss meter. While for the reference sample all reflected rays fall into the receptor field aperture, for the low/mid-gloss sample part of the scattered rays fall out, resulting in a lower flux and gloss value

Table 6: Tolerances on the source and receptor apertures, as defined in ASTM D523-14

	In plane of measurement		Perpendicular to plane of measurement	
	θ (°)	Relative dimension	θ (°)	Relative dimension
Source image	0.75	0.171	2.5	0.568
Tolerance	± 0.25	0.057	± 0.5	0.114
60° Receptor	4.4	1.000	11.7	2.668
Tolerance	± 0.1	0.023	± 0.2	0.046

at the 95% confidence level for reproducibility. These threshold values are also summarized in Table 2. Reproducibility values, i.e., the difference between the maximum and minimum average gloss values obtained with all gloss meters in the appropriate measurement geometry, are presented in Table 4. For the 20° geometry data (samples 15–25), with the exception of the result of sample 15, all values exceed the threshold of 6.4 and 4 GU, respectively. When comparing the results of each pair of gloss meters, reproducibility is obviously better. For three pairs of gloss meters, all reported differences are below the ASTM threshold value of 6.4 GU. The best agreement is found for two gloss meters developed by one and the same manufacturer (with an average difference of 0.8 GU), which is the only pair of gloss meters for which the ISO threshold value of 4 GU is never exceeded. It was expected to find a similar result for the two other manufacturers when comparing their two gloss meters, but this seems not to be the case. The average difference here, respectively, numbers 3.9 and 4.4 GU. Moreover, in these cases, for some samples, the difference is again larger than both threshold values. The reproducibility in the 60° geometry (samples 8–14) is better than in the 20° geometry. Taking into account the results of all five gloss meters, maximum gloss differences are inferior to the thresh-

old values for about half of the samples. For the other samples, gloss differences above threshold are caused by just one gloss meter, for which the results are deviating strongly. Analogous to the 20° geometry, for three gloss meter combinations, all reported differences are below the ASTM threshold of 3.5 GU, while for only two of these three combinations the ISO threshold requirement (3 GU) is met. The best agreement is again found for the same two gloss meters, developed by one and the same manufacturer (with an average difference of 0.6 GU). Finally, in the 85° geometry only the value of sample 7 exceeds the ISO threshold of 2 GU, while all reported values are clearly below the ASTM threshold of 7.2 GU. In this geometry, results of each two gloss meters thus seem to correspond better. The average difference between two instruments' results ranges between 0.5 and 1.3 GU.

In addition to an investigation of the repeatability and reproducibility of the gloss meters, one might also be interested in the correlation between the results of each of the two gloss meters. In Fig. 4, the average gloss value of each sample is presented for each gloss meter, against the average gloss value as obtained with another gloss meter, in the 60° geometry. Each combination of symbol and color represents a different pair of gloss meters. Results obviously suggest a linear

correlation between each of the two gloss meters. As previously stated, the best agreement is obtained for two gloss meters developed by the same manufacturer. The average 60° geometry specular gloss values of each sample obtained with these two devices, denoted as instrument A and instrument B, are plotted against each other in Fig. 5. A similar correspondence between the results of both instruments is found in the 20° and 85° measurement geometry.

For two gloss meters for which the results do not indicate a satisfactory agreement as presented in Fig. 5, a simple correction of the results of one of the two instruments based on the correlation between the two instruments in each measurement geometry might be suggested. By way of example, the usefulness of such a linear correction was evaluated for two gloss meters for which the data show significant differences. For the sake of clarity, these instruments are further denoted as instrument C and instrument D, respectively. In Table 5, the reproducibility as obtained from the difference between the average gloss data of instrument C and instrument D is assessed, before and after a correction based on the calculated correlations. As can be noticed and as expected, the overall average difference decreases, from 4.5 to 3.3 GU. Yet, for 7 resp. 10 samples, the reproducibility thresholds of ASTM and ISO are still not reached. This means that a simple linear correction of the measurement data, based on the correlations between the results of the two gloss meters in each measurement geometry, is not sufficient for effective comparison of the measurement results.

Theoretical approach

Ray tracing simulations

Monte Carlo ray tracing software has been proven to be a useful tool for modeling, analysis, and optimization purposes in multiple applications, such as in computer graphics, luminaire design, and photovoltaics, to mention a few.⁹ In order to get a better idea and understanding of the deviations that might be encountered in specular gloss measurements, ray tracing simulations were put forward.

As a starting point, the optical design of the specular gloss meter as depicted in Fig. 1 was simulated based on the description of the instrumental components and the geometric conditions as stated in ASTM D523-14.³ More specifically, the 60° converging-beam measurement geometry was modeled, as presented in Fig. 6. A light source emits rays of light which are condensed on the source field aperture through a condenser lens. The source field aperture is imaged on the collector lens by use of a projection lens. The receptor field aperture in front of the collector lens restricts the receptor field angle to the values defined in the standard, which are summarized in Table 6 (for the 60° geometry). Through the collector lens, the collected flux finally reaches the photodetector.

To analyze the ray tracing path, each simulation starts from a luminous flux of 1 lm attributed to the source. Next, 1 million rays are traced from the source through the optical system toward the sample position. Rays are reflected from this sample according to the defined scattering behavior or Fresnel reflections. To maximize the amount of rays that reach the detector throughout the optical system, importance sampling^{10,11} from the sample surface toward the detector optical system is applied. As a result of the simulation, the reflected flux captured by the detector is reported.

In order to calculate the specular gloss of the sample under test, two ray tracing simulations have to be performed. A first simulation is performed using the reference sample, which consists of a flat black glass with refractive index of 1.567 at a wavelength of 587.6 nm, and which is attributed a specular gloss value of 100 GU. Afterward, a second simulation is performed at the same wavelength, assigning the surface material reflection properties to the sample under test (see below). Finally, the specular gloss value of the test sample is calculated from the ratio of the reported fluxes.

Bidirectional reflectance distribution function of two test samples

The most universal way to characterize the reflection properties of a material is through the so-called bidirectional reflectance distribution function

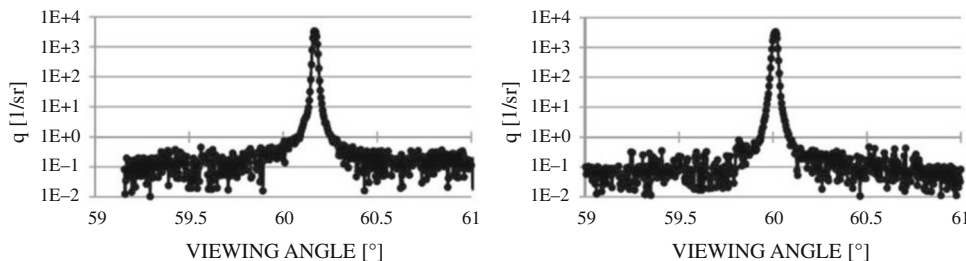


Fig. 7: BRDF of sample 19, measured in-plane (left) and out-of-plane (right) with a viewing angle ranging $\pm 1^\circ$ around the specular reflection direction

Table 7: Influence of admissible tolerances of the source and receptor apertures, as defined in ASTM D523-14, on the reported specular gloss

	Receptor aperture					
	Sample 9 (low/mid-gloss)			Sample 19 (high-gloss)		
	Minimum	Nominal	Maximum	Minimum	Nominal	Maximum
Source aperture						
Minimum	25.7	26.2	26.7	91.0	91.0	91.0
Nominal	25.3	25.8	26.3	91.0	91.0	91.0
Maximum	24.8	25.3	25.8	91.0	91.0	91.0

(BRDF).¹² BRDF is defined as the quotient of the radiance of a surface element in a given direction of view at a given wavelength λ , $dL_r(\theta_i, \theta_r, \lambda)$, by the irradiance incident on the medium from a particular direction of irradiation at the given wavelength, $dE_i(\theta_i, \lambda)$ ²:

$$\rho(\theta_i, \theta_r, \lambda) = \frac{dL_r(\theta_i, \theta_r, \lambda)}{dE_i(\theta_i, \lambda)} \text{ (sr}^{-1}\text{)}. \quad (1)$$

The BRDF can be used to describe the scattering properties of a surface for any combination of illumination and viewing angle.¹³

Two out of the 25 test samples were selected for use in the ray tracing simulations; a white low/mid-gloss sample (sample 9), and a white high-gloss sample (sample 19). For both samples, the BRDF was measured with a goniospectrophotometer developed at LNE-CNAM.¹⁴ The characterization was carried out with a fixed incidence angle of 60°, while the viewing angle was varied around the specular direction with an angular step of 0.004°. Measurements were performed both in the plane of incidence (in-plane) and in a plane perpendicular to the plane of incidence (out-of-plane). As an example, the measurement results for sample 19 are presented in Fig. 7. As expected, a narrow specular peak can be discerned.

Simulation results

The suitability of the simulation method can be checked by calculating the specular gloss value of the two samples from simulations, and comparing the results with the practical measurement results. For sample 9, the simulations report a specular gloss of 25.8 GU, while the average specular gloss from the measurements numbers 27.5 GU. For sample 19, simulations reveal a specular gloss of 91.0 GU, while the average result from the gloss meter measurements numbers 93.4 GU. Taking into account the standard deviation on all experimental measurements, the correspondence between the results can be judged satisfactorily.

The same simulation model can be used to investigate the influence of the admissible tolerances of the

source image and the receptor field aperture (summarized in Table 6) on the gloss measurement result by performing simulations with the upper and lower limit values set to both apertures. Likewise, the influence of a deviation from the ideal angular conditions, i.e., an angular offset on the beam axis angle of $\pm 0.1^\circ$ (e.g., due to a nonflat surface), can be assessed. A summary of the reported specular gloss values for both samples under these conditions is presented in Tables 7 and 8.

As can be concluded from Table 7, the influence of the source and receptor field aperture deviations from nominal values is the highest for the low/mid-gloss sample (sample 9). The largest differences from the specular gloss value for nominal source and receptor field apertures are found for a combination of minimum admissible source field aperture and maximum admissible receptor field aperture, and vice versa. The calculated specular gloss numbers 26.7 and 24.8 GU, respectively. This is a deviation of 0.9 and 1.0 GU from the specular gloss for nominal field aperture values, respectively. For this sample, a maximum deviation of 1.9 GU can be expected taking into account the admissible source and receptor field apertures. The influence of a change in only the source or receptor field aperture is about equal. When changing just one of these, the specular gloss varies about 1.0 GU.

For the high-gloss sample (sample 19), changing the dimensions of the source and receptor field aperture does not seem to have any influence. For each combination of aperture dimensions, the reported specular gloss value is equal to the value found for nominal aperture dimensions, i.e., 91.0 GU. This means that the collected flux is approximately equal for each combination of source and receptor field apertures. Obviously, the reflection properties of the sample play an important role. Indeed, for a high-gloss sample with a distinctive specular peak, such as sample 19, the entire flux comprised in the specular peak region remains within the receptor window, irrespective of the exact aperture dimensions. As such, a departure from the nominal apertures has no influence on the measurement outcome.

Results in Table 8 reveal that the influence of an angular offset on the beam axis angle of 60° with $\pm 0.1^\circ$ has only a minor influence on the reported specular gloss values. Indeed, a difference in specular gloss of

Table 8: Influence of a deviation from the ideal angular conditions, introducing an angular offset of $\pm 0.1^\circ$ to the beam axis angle of 60° , on the reported specular gloss

	Receptor aperture					
	Sample 9 (low/mid-gloss)			Sample 19 (high-gloss)		
	Minimum	Nominal	Maximum	Minimum	Nominal	Maximum
Source aperture						
Minimum	25.6	26.1	26.6	90.7	90.7	90.7
Nominal	25.2	25.8	26.3	90.7	90.7	90.7
Maximum	24.8	25.3	25.8	90.7	90.7	90.7

The admissible tolerances of the source and receptor apertures are again taken into account

only 0.1 to 0.3 GU is found, in comparison to the specular gloss values obtained for the equivalent combination of source and receptor field aperture under ideal angular conditions. Moreover, in accordance with observations from Table 7, about the same maximum differences are found when varying aperture sizes, and an equal influence of a change in only the source or receptor field aperture can be observed.

Discussion and conclusions

When specifying specular gloss meters, manufacturers inform the customer about the quality and accuracy of the device, among others by stating the instrument repeatability and reproducibility. As stated in ISO 2813:2014, the repeatability and reproducibility data reported by manufacturers are typically based on ideal standards.⁴ Therefore, they might be incompatible with the precision data stated in the standards. In practice, there are parameters that can lead to significant higher deviations of measuring values.⁴ Care should therefore be taken by the operator when evaluating gloss measurement results.

Experimental specular gloss measurement results obtained with six gloss meters and 25 samples confirm this notion. Neglecting the results of the four samples with a restricted gloss uniformity, average repeatability values range between 0.4 and 8.3 GU. In general, these values are higher than the threshold values defined in the ASTM standard. One explanation could be the relative importance of outliers. However, even neglecting the measurement results which deviate more than 5 GU in each series of 5 measurements for the 21 test samples that are free from surface directionality,⁴ average repeatability values still remain above threshold for the same test samples as reported in Table 3, and range now between 0.4 and 4.4 GU.

For reproducibility, differences of up to 10 GU were found. Yet, differences depend on the measurement geometry; a better reproducibility is obtained for the 85° and 60° geometry, compared to the 20° geometry. Based on this finding, it could be questioned if stating only one repeatability and reproducibility value for all

three measurement geometries and for the entire gloss measurement range is adequate. Furthermore, reproducibility seems to depend on which pair of gloss meters is being compared. Although, as expected, the best agreement is obtained for two test instruments developed by the same manufacturer, this statement cannot be generalized. In the comparison between the two other manufacturers' gloss meters, important differences are recorded. Finally, while previous research has questioned the linearity between gloss measurement results obtained with different instrumentation,⁷ from the present data, a linear correlation between the results of each two gloss meters can be put forward. A correction of the measurement results based on the correlation found, however, does not always guarantee that reproducibility values get below threshold.

Since commercial gloss meters have compact sizes, the distances between the elements and apertures are small. This may easily lead to errors in the setting of the specular angle, measurement of source and receptor apertures, and positioning of the focused image of the source in the center of the receptor entrance window. In trying to elucidate some of the reported differences, both the influence of the admissible tolerances on the source and receptor field aperture, and of a deviation from the theoretic 60° measurement geometry, was investigated by use of ray tracing simulations on two test samples. Simulations indicate that differences in aperture dimensions may lead to gloss differences of about 2 GU in the case of the low/mid-gloss sample. This is in line with the statement made in ASTM D 523-14, which, based on a study performed by Hammond and Nimeroff,¹⁵ indicates that tolerances are chosen such that errors in the source and receptor apertures do not produce an error of more than ± 1 GU at any point on the scale.³ In line with Hammond and Nimeroff,¹⁵ the influence of errors in the source and receptor apertures is found to be more important than the influence of an angular offset on the beam axis angle. According to Hammond and Nimeroff, it should be noted that errors of 5 GU might still be encountered if all variables happen to contribute errors in the same direction.¹⁵

Finally, some other sources of error can be identified. Deviations from the defined spectral conditions (spectral flux distribution equal to CIE illuminant C, and spectral response of the receiver identical to the spectral luminous efficiency function $V(\lambda)$) may lead to higher measurement errors for chromatic, low-gloss samples. Also the degree of polarization and the stability of the illumination beam might have a substantial effect. It is not clear to which extent these potential sources of error are taken into account by the instrument manufacturers. Reference goniophotometers for specular gloss measurements with relative expanded uncertainties of about 0.4% for each of the three measurement geometries have been reported.¹⁶ Taking into account that the practical primary gloss standards also deviate from their theoretical counterpart,¹⁷ comparison of measurements performed at national metrology institutes indicate measurement uncertainties of up to 0.6 GU for high-gloss working samples.¹⁸ It is therefore clear that gloss measurements using commercial gloss meters may engender measurement uncertainties far beyond 1 GU. More information on how the repeatability and reproducibility of commercial gloss meters is being determined should therefore be made public, while repeatability and reproducibility values should be stated for each measurement geometry and gloss range. In this respect, further ray tracing simulations using a variety of BRDF distributions in the different gloss measurement geometries could provide useful information.

Acknowledgments This work has been partially funded by the European Metrology Research Programme (EMRP). The EMRP is jointly funded by the European Commission and participating countries within the European Association of National Metrology Institutes (EURAMET).

References

1. CIE Publication 175:2006, “A Framework for the Measurement of Visual Appearance.” CIE, Vienna (2006)
2. CIE Publication S 017/E:2011, “International Lighting Vocabulary.” CIE, Vienna (2011)
3. ASTM Standard D 523-14, “Standard Test Method for Specular Gloss.” In: *Annual Book of ASTM Standards*, vol. 06.01. ASTM International, West Conshohocken, PA (2015)
4. ISO Standard 2813:2014, “Paints and varnishes – Determination of gloss values at 20 degrees, 60 degrees and 85 degrees.” International Organization for Standardization, Geneva, Switzerland (2014)
5. Cook, MP, Thomas, K, “Evaluation of Gloss Meters for Measurement of Moulded Plastics.” *Polym. Test.*, **9** (4) 233–244 (1990)
6. Sève, R, “Problems Connected with the Concept of Gloss.” *Color Res. Appl.*, **18** (4) 241–252 (1993)
7. Arney, JS, Ye, L, Banach, S, “Interpretation of Gloss Meter Measurements.” *J. Imaging Sci. Technol.*, **50** (6) 567–571 (2006)
8. TAPPI Test Method T 480, “Specular Gloss of Paper and Paperboard at 75 Degrees.” TAPPI, Peachtree Corners, GA (2009)
9. Leyre, S, Ryckaert, J, Acuña, P, Audenaert, J, Meuret, Y, Durinck, G, Hofkens, J, Deconinck, G, Hanselaer, P, “A Hybrid Tool for Spectral Ray Tracing Simulations of Luminescent Cascade Systems.” *Opt. Express*, **22** (20) 24582–24593 (2014)
10. Veach, E, “Robust Monte Carlo Methods for Light Transport Simulation.” Doctoral Dissertation. Stanford University, Stanford, CA (1998)
11. Dutré, P, Bala, K, Bekaert, P, *Advanced Global Illumination*, 2nd ed. A K Peters Ltd, Wellesley, MA (2006)
12. Nicodemus, FE, Richmond, JC, Hsia, JJ, Ginsberg, IW, Limperis, T, “Geometric Considerations and Nomenclature for Reflectance.” *National Bureau of Standards*, Monograph 160 (1977)
13. Leloup, FB, Forment, S, Dutré, P, Pointer, MR, Hanselaer, P, “Design of an Instrument for Measuring the Spectral Bidirectional Scatter Distribution Function.” *Appl. Opt.*, **47** (29) 5454–5467 (2008)
14. Obein, G, Audenaert, J, Ged, G, Leloup, FB, “Metrological Issues Related to BRDF Measurements Around the Specular Direction in the Particular Case of Glossy Surfaces.” *Proc. SPIE*, 9398, 93980D, San Francisco, March 2015
15. Hammond, HK, III, Nimeroff, I, “Measurement of Sixty-Degree Specular Gloss.” *J. Res. NBS*, **44** 585–598 (1950)
16. Nadal, ME, Thompson, EA, “NIST Reference Goniophotometer for Specular Gloss Measurements.” *J. Coat. Technol.*, **73** (917) 73–80 (2001)
17. Nadal, ME, Thompson, EA, “New Primary Standard for Specular Gloss Measurements.” *J. Coat. Technol.*, **72** (911) 61–66 (2000)
18. Nadal, ME, Zwinckels, JC, Noël, M, “Specular Gloss Scales Comparison Between the National Institute of Standards and Technology and the National Research Council of Canada.” *J. Coat. Technol.*, **75** (943) 45–51 (2003)