

Stability of Ceramic Color Reflectance Standards

Hugh S. Fairman,* Henry Hemmendinger

Hemmendinger Color Laboratory, 438 Wendover Drive, Princeton, New Jersey 08540

Received 11 August 1997; 13 May 1998

Abstract: This article reviews the long-term repeatability of spectrophotometric color measurements of ceramic color standards calibrated by Hemmendinger Color Laboratory (HCL) since it was founded 25 years ago. A set of twelve BCRA tiles was measured at HCL in 1977, and a few months later at NBS. The CIELAB color-difference, averaged over the twelve tiles, between HCL and NBS was 0.25 units. Measured at HCL about every four months since that time, the repeatability of measurement, averaged over the twelve tiles, has been better than 0.15 CIELAB units. The difference from the NBS measurements has remained at about 0.25 units over two decades. From these data the conclusion can be drawn that the tiles have not changed color during this period of time by an amount exceeding the limits of instrumental repeatability, namely about 0.15 units. © 1998 John Wiley & Sons, Inc. Col Res Appl, 23, 408–415, 1998

Key words: calibration; ceramic tiles; spectrophotometry

INTRODUCTION

About thirty years ago, the British Ceramic Research Association and the National Physical Laboratory of Great Britain initiated a program to provide sets of ceramic tiles for use as spectrophotometrically and colorimetrically defined color standards. Extensive studies^{1,2} preceded the selection of appropriate formulations, culminating in the selection of a set of twelve tiles, designated *British Ceramic Colour Standards, Series I*; the first shipments were made in 1969. After about a decade, a review of their use led to the formulation of a new set of twelve colors, designated *Ceramic Colour Standards Series II*.³ These standards have been well accepted and widely used throughout the world. A series of articles by NPL personnel has described their use.^{1–4}

Hemmendinger Color Lab (HCL) was established in 1970 by one of us (HH). A major objective of the new laboratory was the determination of the repeatability, and

ultimately the accuracy, with which commercially available instruments could be used over an extended period of time. This article reports on the experiences over more than two decades. Although related data have been published in the reports of national standards laboratories, we believe that this is the first report summarizing the results obtained over twenty years with commercially available instruments and standards.

For twenty-five years the Hemmendinger Color Lab (HCL) has monitored its performance by measurement of a set of British Ceramic Research Association (BCRA) tiles. This set of tiles is from Series I, and carries the serial number HCL-4. From 1972–1992, measurements were made on a GE–Hardy Spectrophotometer, modified to provide high-resolution (0.01%) digital output. After 1992, measurements were made on a Datacolor Spectraflash 500 instrument with 0.001% resolution.

The Hardy spectrophotometer was calibrated with white reflectance standards and colored filters obtained from the National Bureau of Standards (NBS now NIST). As NBS procedures were improved, new NBS standards were introduced at HCL. In 1983, a change was made in the reduction of the Hardy data: an empirical correction was made in the Hardy photometric scale in recognition of a systematic difference of about 0.3% (maximum) that had existed between NBS and HCL Hardy data.

There are three principal periods of the laboratory's operations. The Hardy was used from 1972–1983 using NBS reference standards, by optimizing the numerous adjustments available on the Hardy instrument, and with no data-reduction corrections. From 1983–1993, the Hardy data were modified to fit the photometric scale defined by the three gray tiles of BCRA Series I. Since 1993, the Spectraflash 500 was used based on a photometric scale from National Physical Laboratory (NPL), Teddington, England and with wavelength calibration based on laser emissions.

This article reviews the accuracy and repeatability with which BCRA set HCL-4 has been measured. Over each of these three periods, the mean CIELAB color-difference

*Correspondence to: Mr. Hugh S. Fairman
© 1998 John Wiley & Sons, Inc.

TABLE I. Color names, luminance, and chromaticity coordinates of Set HCL-4, BCRA Ceramic Tiles, Series I.

Sample No.	Name	Y	x	y
1	Light grey	62.3	0.315	0.337
2	Medium grey	29.2	0.306	0.329
3	Dark grey	9.27	0.306	0.322
4	Maroon	6.41	0.373	0.321
5	Pink	44.3	0.363	0.327
6	Brown	20.0	0.450	0.380
7	Yellow	64.7	0.419	0.459
8	Light green	34.3	0.310	0.372
9	Dark green	10.8	0.316	0.401
10	Greenish blue	11.4	0.264	0.325
11	Medium blue	16.0	0.257	0.275
12	Dark blue	5.56	0.270	0.254

between the NBS calibration and HCL has been 0.2–0.3 averaged over the twelve tiles.

HISTORICAL SUMMARY OF MEASUREMENTS

Samples and Instruments

Years 1972–1982

The first set of British Ceramic Research Association (BCRA) standards obtained in 1972 was measured on a General Electric–Hardy spectrophotometer, serial number 4092225. Its calibration was carried out according to procedures⁵ based on standards from the US National Bureau of Standards. The photometric standards available from NBS for the first work at HCL were white Vitrolite reference standards. For wavelength accuracy, the primary reference from NBS at that time was didymium glass, though it was recognized that the use of didymium for assessment of wavelength accuracy had serious limitations resulting from the interaction between wavelength variations and band-width changes. In 1965, NBS issued a series of glass filters^{6,7} for the evaluation of the then newly developed spectrophotometer-integrator systems; these filters were offered as SRM 2101-2105 without spectrophotometric data, but with colorimetric specifications only. With the aid, however, of publication by NBS⁸ of spectral data for the master set (retained at NBS) of the SRM 2101-2105 series, we were able to derive spectral data consistent with the NBS scale. In 1975, after the construction of the 0.01% Reference Spectrophotometer^{9–11} at NBS, these spectral data were

supplemented by high-precision NBS measurements of some of the filters in the SRM 2101 series. This first period in the operation of HCL, with the General Electric spectrophotometer and with traceability to NBS standards, extended from 1972–1982. However, since the first half of this period, from 1972–1977, was based on improvised standards, measurements over that period are not included in the present recapitulation.

Years 1983–1993

In 1983, the HCL laboratory was moved from Belvidere, NJ to Princeton, NJ. By then we had recognized that there was a significant error (at its maximum in mid-scale, about 0.3%) in the then currently used photometric scale of the GE instrument. Correction to the NBS reflectance scaling was, therefore, made on the basis of the NBS measurements of the three grays in the set of BCRA standards used at HCL. The second period of measurement-calibrations was from 1983–1993; it differed from the previous period only in this correction of photometric scale based on NBS measurements of the three grays of BCRA Series I tiles.

Years 1993–1997

By 1993, a newly designed spectrophotometer was available from Instrumental Colour Systems (now part of Datacolor), and it became evident that the long life of the General Electric instrument as the paramount one for reflectance color measurements had ended. Since 1994, measurements at HCL have been based on the Datacolor Spectraflash 500, serial number 0296 CT.

The continuity of measured data has required the continued use of the Series I color standard initiated in 1972. Since these have now become obsolete, we describe the colors used in this study in Table I, measured with specular component included.

Criteria for Evaluation of Samples and Instruments

To judge the durability of color standards, it is necessary to use measurements; the measurements must be made with instruments of known accuracy, and the accuracy of an instrument must be evaluated with standards of known reflectance. It is accordingly essential that the standards used to evaluate instrument accuracy or repeatability over

TABLE II. The periods of the measurements as defined by the instrument and conditions of calibration.

Period	Dates	Instrument and calibration
P1	1978–1982	General Electric–Hardy, Serial No. 4092225, fitted with digital readout of 0.01% resolution, and based on NBS absolute white scale of 1978.
P2	1983–1993	As P1, but with application of a photometric scale correction based on NBS grey tile measurements of 1979, and relative to the 1981 NBS absolute reflector.
P3	1993–1997	Datacolor Spectraflash 500, Serial No. 0296 CT, calibrated by the manufacturer to agree with the NPL (National Physical Laboratory, Great Britain) absolute reflector of 1989.

BCRA Series I
Effect of $\Delta X = +0.02$ and $\Delta Y = -0.02$

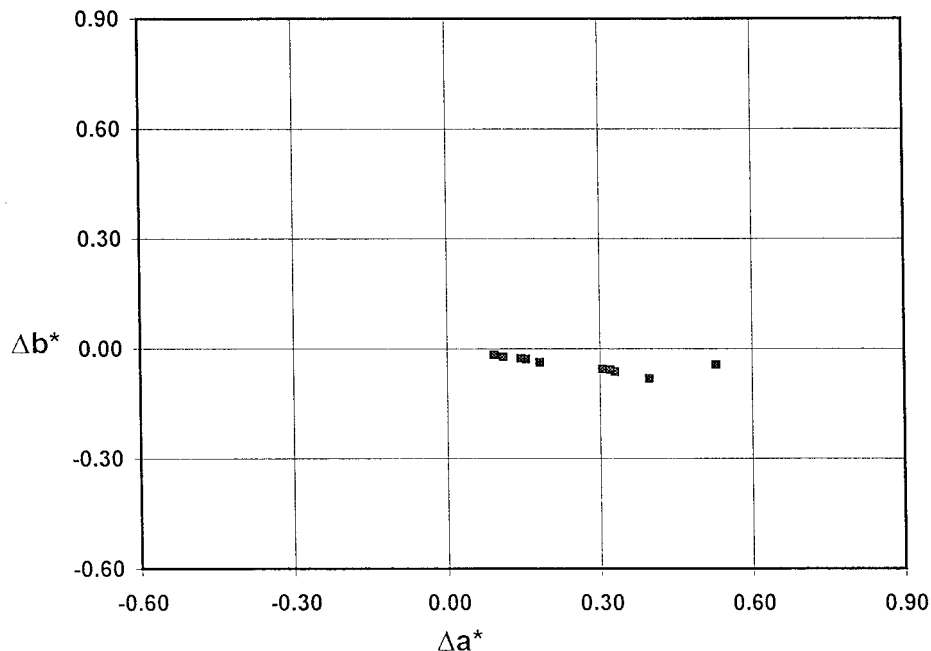


FIG. 1. The colorimetric effect in Δa^* and Δb^* of a difference of $+0.02$ in the X tristimulus value and -0.02 in the Y tristimulus value from the standard tristimulus values of the 12 BCRA Series II Ceramic Colour Tiles. The Z value is held constant and the Δa^* , Δb^* differences calculated. One point is off the scale of the diagram, but the scale is consistent with the scale of all other Δa^* and Δb^* plots in this article.

an extended period be different from those used to determine the durability of the samples. In the view of the authors, there is no other set of standards today that can rival the BCRA Series II. Until we have redundancy in the sets of available standards, this leaves us in a “catch 22” situation. The situation is somewhat alleviated by the use not only of the scalar quantity ΔE^* alone, but also of its three component vectors.

There are, however, some situations in which the use of the scalar ΔE^* is necessary. Consider the question, “Can one make any useful generalization about the magnitude of the change in reflectance, which generates a given magnitude of color difference?” From considerations described later in this article, we conclude that spectral differences with an amplitude of 0.2% will frequently correspond to more than 0.4 CIELAB unit. Standards of this 0.2% reflectance accuracy are not generally available, and we must recognize that any consideration of the durability of laboratory standards must be limited to the accuracy of the available reference standards. To some extent, it is, however, the reproducibility of the laboratory measurements that is primarily at the heart of the evaluation of durability. In the following report, we accordingly emphasize the reproducibility of measurements over 25 years, and our conclusions are further reinforced by the fact that two commercially manufactured instruments have been used, and they agree with one another with a precision as good as, or better than, the guaranteed accuracy of national laboratory standards.

All data reported here are the averages of four measure-

ments made on four days, normally at least a week apart. The repeatability of the instrument used in Period 3 leads to standard deviations of 0.01 or 0.02%, except in spectral regions where $dR/d\lambda$ becomes large.

EXPERIMENTAL DATA

Measurements were made on a set of BCRA (Series I) Tiles, Serial Number HCL-4. The tiles were usually measured several times each year since 1972. The data are reported in three groups, as shown in Table II. The three groups repre-

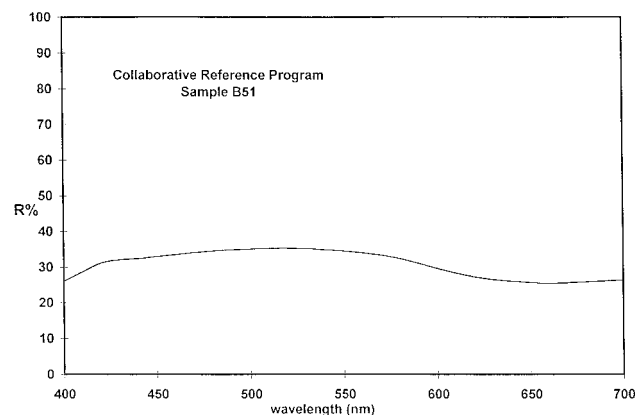


FIG. 2. The reflectance spectrum of a specimen from the Collaborative Reference Program of about two years ago. This specimen is the reference specimen for the data of Figs. 3 and 4.

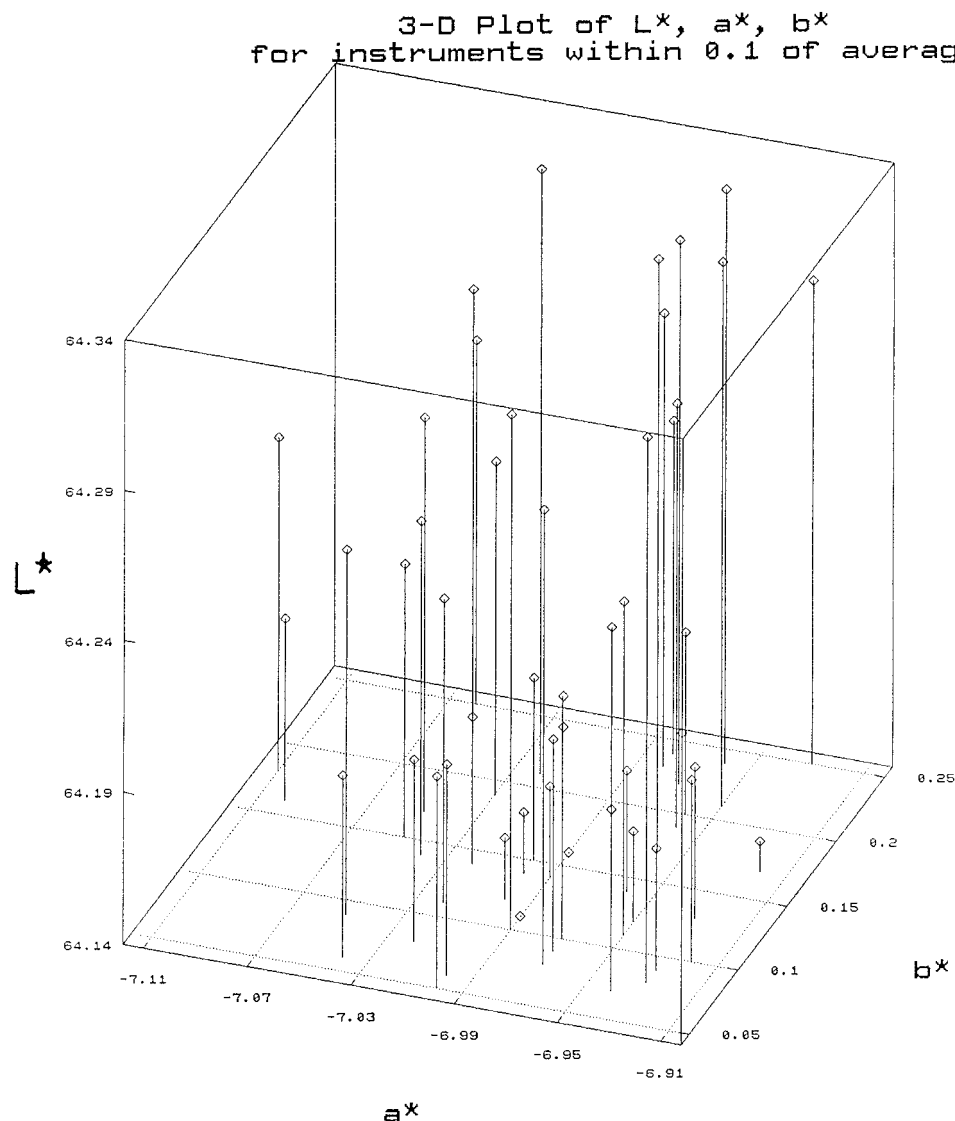


FIG. 3. Forty-seven of approximately 440 instruments measuring the specimen of Fig. 1 report its colorimetric values within 0.1 CIELAB unit of the average of all instruments. The reported values are plotted on a Δa^* , Δb^* , and ΔL^* 3-D diagram whose axes extend ± 0.1 unit from the average of all 440 instruments.

sent periods where the laboratory's instrument or method of calibration differed from period to period.

Graphical Presentation of Data

The data are too extensive for presentation in tabular form. They are, therefore, presented in a series of graphs that are representative of the complete data. However, before showing this sampling, it is useful to take two steps: to consider how many components of the vector ΔE^* can be effectively displayed, and to consider the significance of small differences in the measured data.

To meet the objective of this article, we should display measurements of twelve samples over a period of several decades. We have found no way to display effectively the three dimensions required. We, therefore, look only at the

two chromaticity dimensions Δa^* and Δb^* . The variance in ΔL^* was usually no larger than that of either Δa^* or Δb^* .

Scale for Graphical Presentation of Data

The data are reported in terms of colorimetric differences. The magnitudes of the differences are highly dependent on the shape of the color-matching functions and the spectral curve of the measured specimen. Therefore, it seems likely that the relationship of spectral difference in percent reflectance can only be anecdotally related to colorimetric difference, and we propose to do that through the use of the following three scenarios.

CRP 92
Delta R% from mean of all measurements

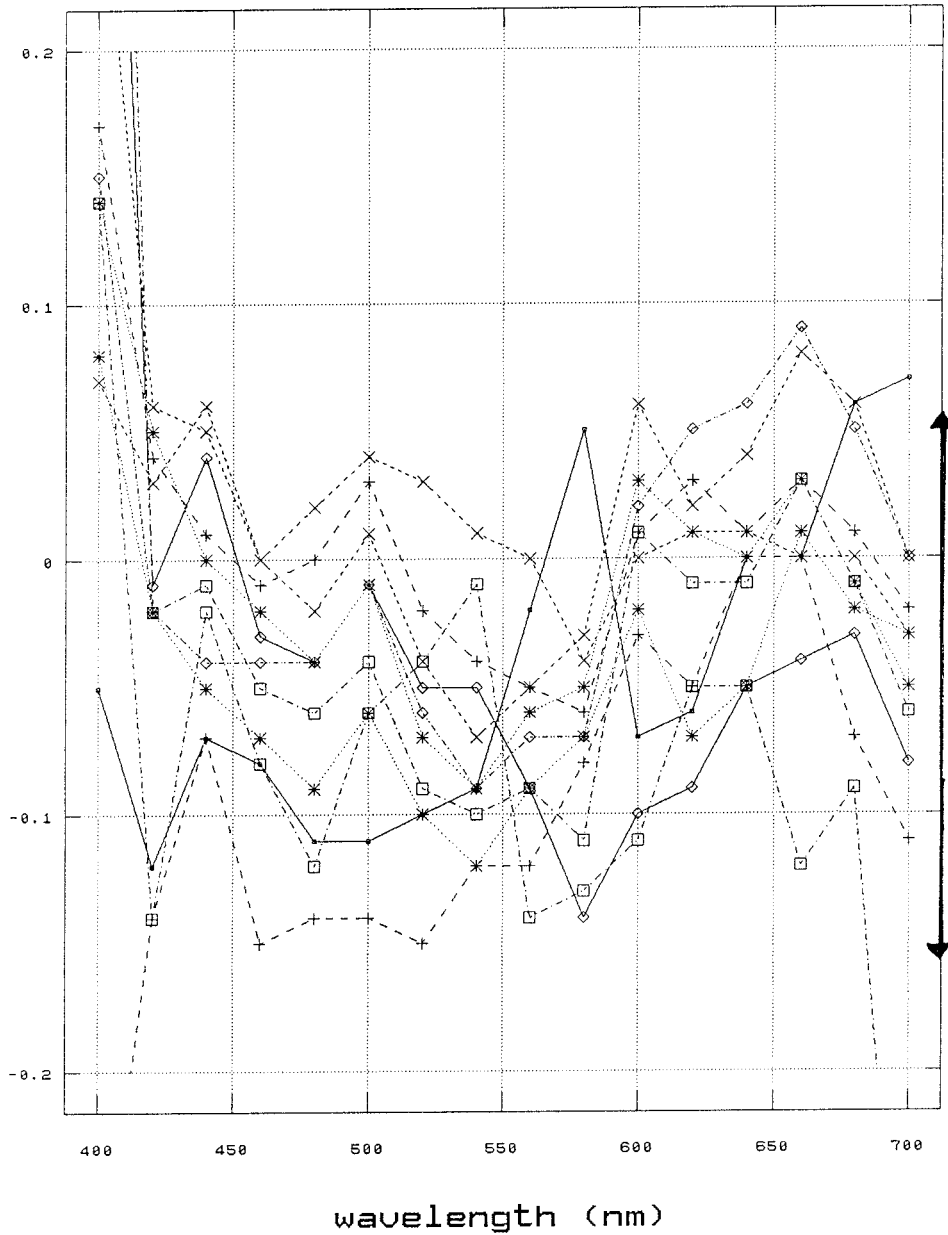


FIG. 4. Eleven of the 47 instruments plotted in Fig. 3 participated in another experiment where spectral reflectances for this specimen on these instruments are obtained. The reflectances are plotted by difference from the average of all instruments participating. Reflectance differences of about ± 0.1 , as shown by the bold-face arrow, lead to about ± 0.2 CIELAB unit differences for this specimen.

Scenario 1

Figure 1 shows the colorimetric effect in a Δa^* and Δb^* diagram of a change of $+0.02$ in X value and -0.02 in Y value from the standard values of each of the twelve tiles of BCRA Series I Ceramic Colour standards. The Z tristimulus value has been held constant. All data referred to in Fig. 1 and subsequent figures are calculated using CIE Standard Illuminant D65 and the CIE 1931 Standard Observer. The chromaticness differences that result from these differences in tristimulus value for each tile are plotted on the same

scale as that which will be used for all subsequent figures displaying scatter. Note that one point is actually off the scale on this diagram, but for ease of interpretation one scale is used throughout.

The distribution of chromaticness differences range over nearly one CIELAB unit of color difference. These are representative of the colorimetric difference introduced by this magnitude of tristimulus difference. The direction and apparent linearity are probably not significant. Tristimulus differences were chosen because major-axes of MacAdam ellipses generally lie around a mean of about 45° in the CIE

TABLE III. The colorimetric worst-case effect of a difference of 0.2% reflectance on selected BCRA Series II ceramic tiles. All data expressed as ΔE^*_{ab} for Standard Illuminant D65 and the 1964 10° Standard Observer.

Tile color	Mode of measurement	
	Specular included	Specular excluded
Deep blue	1.9	5.1
Green	0.9	1.1
Bright yellow	0.7	0.8
Orange	0.8	1.2
Red	1.2	2.7

xy chromaticity diagram. Thus, the chosen differences generally move along the semi-minor axes and will tend to define the largest ratio of difference to unit change in tristimulus value in a more uniform color space such as CIELAB.

Scenario 2

Figure 2 shows a reflectance spectrum of a specimen measured in the Collaborative Reference Program (CRP) about two years ago. Nearly 440 instruments reported CIELAB notations on this specimen. Although a sphere of three units diameter in CIELAB space would be required to fully encompass all the reported data on this one specimen, in Fig. 3 we have restricted the volume to about 0.2 CIELAB units (0.1 plus and 0.1 minus from the mean of all data) in each of the three directions. This includes the data of 47 instruments. Of the 47 instruments, eleven partici-

pated in another experiment where their spectral data was also included. Thus, it is possible to plot the spectral data, also by difference from the mean of all spectral data on this specimen. Figure 4 suggests that an instrument must be within about $\pm 0.1\%$ reflectance to be within 0.2 CIELAB units in color difference for this arbitrary specimen. Similar results have been found in the analysis of earlier CRP reports.

Scenario 3

Table III illustrates the color difference introduced by 0.2% reflectance uncertainty imposed upon the spectra of some selected BCRA Series II tiles for some different instrument geometries and modes of view. The uncertainties are deliberately imposed upon the spectra to provide worst-case effects. The 0.2% reflectance was chosen for this illustration, because the stated uncertainty of any national standardizing laboratory is never less than this amount for reflectance mode measurements of chromatic specimens. Indeed, that uncertainty, when stated by the NIST, implies a coverage factor of only one; that is, what most would consider a one σ uncertainty.

TYPICAL RESULTS

In the following three figures, the scatter of individual tiles over the approximately 21-year period, which encompasses the three defined periods of differing calibration methods, are plotted one tile at a time. The three periods are keyed by individual symbols for each period. The three tiles are selected as representative of the rest of the data. In Fig. 5,

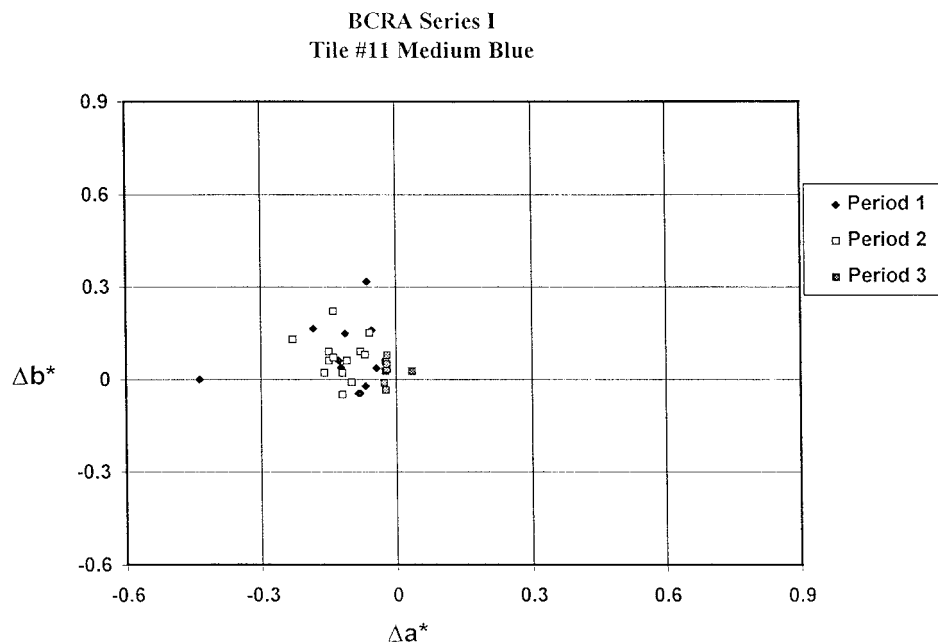


FIG. 5. Colorimetric results from calibrations of the Medium Blue Tile over the three periods described in the text. This and the following two figures are plotted on a Δa^* and Δb^* diagram, called a "chromaticness diagram." The standard is the NBS 1978 measurement of the Medium Blue Tile. The calibrations improve for both accuracy and precision over time. This tile is typical of the performance of most tiles. Six other tiles exhibit this improved performance over time.

BCRA Series I
Tile #7 Yellow

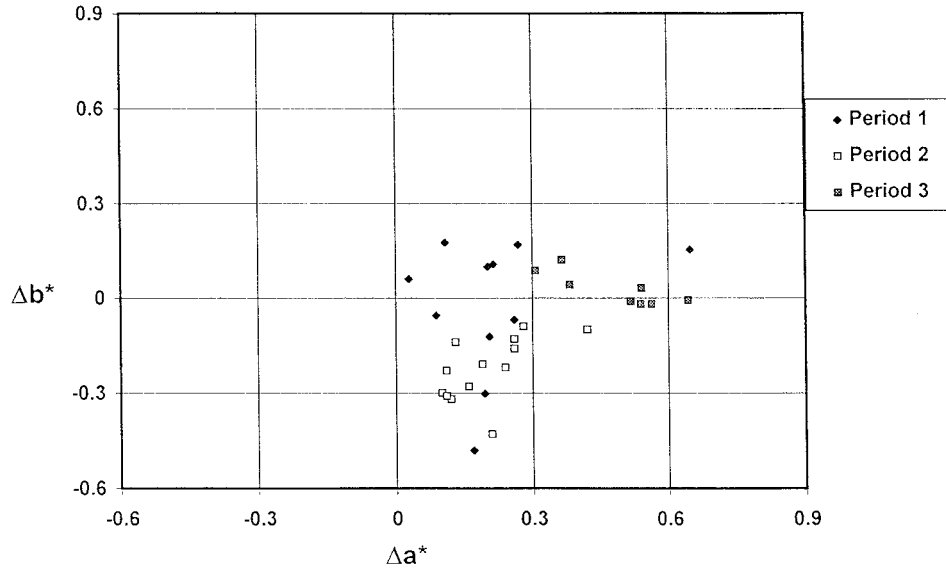


FIG. 6. Colorimetric results from calibrations of the Yellow Tile over the three periods described in the text. The calibrations improve for precision over the three time periods, but accuracy deteriorates in the third period.

the chromas of the #11 Medium Blue tile from eleven calibrations from Period I, thirteen calibrations from Period II, and eight calibrations from Period III are presented. One can see from this plot that both the accuracy and the precision of the results are improving over time.

Accuracy is defined as the degree to which the results agree with a known value (here, as throughout, the NIST measurement of this tile in 1978). Precision is defined as the closeness of agreement between independent measurement

results (here, as throughout, the several calibrations of a tile within any of the three periods). The results for this Medium Blue tile are presented as representative of the majority of the results obtained in this study with about six more tiles showing performance where both the accuracy and the precision improved over time.

Figure 6 presents the results of the same number of calibrations over the same three periods for the #7 Yellow Tile. This tile is of interest because it is identical in color to

BCRA Series I
Tile #9 Dark Green

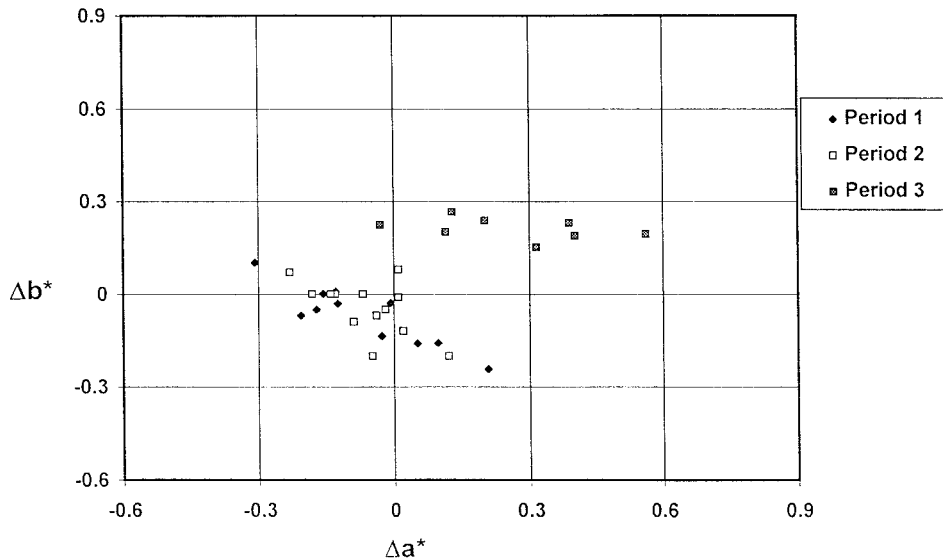


FIG. 7. Colorimetric results from calibrations of the Dark Green Tile over the three periods described in the text. The calibrations retain their level of precision over the three time periods, but accuracy deteriorates over time.

the Bright Yellow tile of Series II, and because it is highly thermochromic.¹² It should be mentioned here that all measurements made in these calibrations are made at temperatures controlled to within 1°C of 25°C. The mean temperature during the four measurements is within 0.2°C of that standard temperature. Since the effects of thermochromism on reflectance are linear over small temperature differences, this averaging of measurement temperature of the tile eliminates temperature difference as a significant source of error.

In Fig. 6, we see that the dispersion of the results, the precision, improves with each succeeding period, but that the absolute difference as represented by the difference between the NBS measurements and the HCL measurements is actually greater in the last period than in the first two periods.

Figure 7 presents the results of Tile #9 Dark Green over the same number of calibrations over the same periods. In this figure, while Periods I and II are quite similar to each other for both accuracy and precision, Period III shows deterioration of both accuracy and precision.

CONCLUSIONS

We have demonstrated the ability to measure a single set of permanent colored materials over a long period of time, by using two instruments and three traceability schemes, with a high degree of precision and accuracy. In fact, it may be seen that no improvement can be made beyond this level of precision or accuracy, because the present values of these parameters are well within the stated uncertainty of a calibration from the national standardizing laboratory.

We may be practically assured that the colors of the set of tiles have not changed over the more than 20-year period encompassed by this study—at least within the limits of our ability to measure such change.

We conclude that the tools, techniques, material standards, and instruments are in place such that we may expect the uncertainty of measurement, not the total uncertainty, in a

carefully operated laboratory in the field to be as low as 0.05–0.1% reflectance. That leads us to conclude that the national standardizing laboratory ought to establish a priority to reduce their uncertainty to at least this level, if not beyond.

1. F. J. J. Clarke and P. R. Samways, The spectrophotometric properties of a selection of ceramic tiles, *National Physical Laboratory Report MC2*, August 1968, London.
2. F. J. J. Clarke, Problems of spectrofluorimetric standards for reflection and colorimetric use, *National Physical Laboratory Report MOM 12*, August 1968, London.
3. J. F. Verrill and F. Malkin, Thermochromic behavior of the ceramic colour standards, in ISCC 61st Annual Meeting and AIC 1992 Meeting, Princeton, New Jersey, June 21–22, 1992.
4. F. Malkin, The Ceramic colour standards, *British Ceramic Research Association Research Paper 702*, 1979.
5. Henry Hemmendinger, The calibration of a spectrophotometer for color measurement, *Color Technology in the Textile Industry*, G. Celikiz and R. Kuehni, Eds., American Association of Textile Chemists and Colorists, Research Triangle Park, NC, 1983.
6. William H. Venable, Jr. and Kenneth L. Eckerle, Didymium glass filters for calibrating the wavelength scale of spectrophotometers—SRM 2009, 2010, 2013, and 2014, *NBS Special Publication 260-66*, U.S. Government Printing Office, Washington, D.C., 1979.
7. Harry J. Keegan, John C. Schleiter, and Deane B. Judd, Glass filters for checking performance of spectrophotometer-integrator systems of color measurement. *J. of Research of the Nat. Bur. Stds.* **66A**, 203–221 (1962).
8. Kenneth L. Eckerle and William H. Venable, Jr., 1976 remeasurement of NBS spectrophotometric-integrator filters. *Color Res. Appl.* **9**, 15–22 (1984).
9. K. D. Mielenz and K. L. Eckerle, Design, construction, and testing of a new high accuracy spectrophotometer, *NBS Technical Note 729*, National Bureau of Standards, June 1972.
10. Kenneth L. Eckerle, Modification of an NBS reference spectrophotometer, *NBS Technical Note 913*, National Bureau of Standards, July 1976.
11. William H. Venable, Jr., Jack J. Hsia, and Victor R. Weidner, Development of an NBS reference spectrophotometer for diffuse transmittance and reflectance, *NBS Technical Note 594-11*, National Bureau of Standards, October 1976.
12. J. Anne Compton, The thermochromic properties of the ceramic colour standards. *Color Res. Appl.* **9**, 15–22 (1984).