Cross-cultural variation of memory colors of familiar objects

Kevin A.G. Smet,[1,*] Yandan Lin,[2] Balázs V. Nagy,[3,4] Zoltan Németh,[3] Gloria L. Duque-Chica,[4] Jesús M. Quintero,[5] Hung-Shing Chen,[6] Ronnier M. Luo,[6,7] Mahdi Safi,[8] and Peter Hanselaer[1]

[1]*ESAT/Light and Lighting Laboratory, KU Leuven, Ghent, Belgium*
[2]*Institute for Electric Light Sources, Fudan University, Shanghai, China*
[3]*Department of Mechatronics, Optics and Engineering Informatics, Budapest University of Technology and Economics, Budapest, Hungary*
[4]*Center for Neuroscience and Behavior, and Department of Experimental Psychology, Institute of Psychology, University of São Paulo, São Paulo, Brazil*
[5]*Department of Electrical and Electronics Engineering, Universidad Nacional de Colombia, Bogotá, Colombia*
[6]*Graduate Institute of Electro-Optical Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan*
[7]*School of Design, University of Leeds, Leeds, UK*
[8]*Color Imaging and Color Image Processing Department, Institute for Color science and Technology, Tehran, Iran*
[*]*Kevin.Smet@kuleuven.be*

Abstract: The effect of cross-regional or cross-cultural differences on color appearance ratings and memory colors of familiar objects was investigated in seven different countries/regions – Belgium, Hungary, Brazil, Colombia, Taiwan, China and Iran. In each region the familiar objects were presented on a calibrated monitor in over 100 different colors to a test panel of observers that were asked to rate the similarity of the presented object color with respect to what they thought the object looks like in reality (memory color). For each object and region the mean observer ratings were modeled by a bivariate Gaussian function. A statistical analysis showed significant ($p < 0.001$) differences between the region average observers and the global average observer obtained by pooling the data from all regions. However, the effect size of geographical region or culture was found to be small. In fact, the differences between the region average observers and the global average observer were found to of the same magnitude or smaller than the typical within region inter-observer variability. Thus, although statistical differences in color appearance ratings and memory between regions were found, regional impact is not likely to be of practical importance.

OCIS codes: (330.1690) Color; (330.1720) Color vision; (330.5510) Psychophysics; (330.4060) Vision modeling. (330.5020) Perception psychology

References and links

1. Hering, *Grundzüge der Lehre vom Lichtsinn* (Springer-Verlag, 1920).
2. C. J. Bartleson, "Memory colors of familiar objects," J. Opt. Soc. Am. **50**(1), 73–77 (1960).
3. S. M. Newhall, R. W. Burnham, and J. R. Clark, "Comparison of successive with simultaneous color matching," J. Opt. Soc. Am. **47**(1), 43–54 (1957).
4. K. A. G. Smet, W. R. Ryckaert, M. R. Pointer, G. Deconinck, and P. Hanselaer, "Color appearance rating of familiar real objects," Color Res. Appl. **36**(3), 192–200 (2011).
5. M. Vurro, Y. Z. Ling, and A. C. Hurlbert, "Memory color of natural familiar objects: Effects of surface texture and 3-D shape," J. Vis. **13**(7), 20 (2013).
6. P. Siple and R. M. Springer, "Memory and preference for the colors of objects," Percept. Psychophys. **34**(4), 363–370 (1983).
7. J. Pérez-Carpinell, M. D. de Fez, R. Baldoví, and J. C. Soriano, "Familiar objects and memory color," Color Res. Appl. **23**, 416–427 (1998).
8. S. N. Yendrikhovskij, F. J. J. Blommaert, and H. de Ridder, "Representation of memory prototype for an object color," Color Res. Appl. **24**(6), 393–410 (1999).

Received 14 Oct 2014; revised 3 Dec 2014; accepted 3 Dec 2014; published 23 Dec 2014
29 Dec 2014 | Vol. 22, No. 26 | DOI:10.1364/OE.22.032308 | OPTICS EXPRESS 32308

9. C. J. Bartleson, “Color in memory in relation to photographic reproduction,” Photon. Sci. Eng. **5**, 327–331 (1961).
10. C. L. Sanders, “Color preferences for natural objects,” J. Illum. Eng. **54**, 452–456 (1959).
11. D. B. Judd, “A flattery index for artificial illuminants,” J. Illum. Eng. **62**, 593–598 (1967).
12. W. A. Thornton, “A Validation of the Color-Preference Index,” J. Illum. Eng. **4**(1), 48–52 (1974).
13. C. J. Bartleson and C. P. Bray, “On the preferred reproduction of flesh, blue-sky, and green-grass colors,” Photon. Sci. Eng. **6**, 19–25 (1962).
14. J. J. M. Granzier and K. R. Gegenfurtner, “Effects of memory color on color constancy for unknown colored objects,” J. Illum. Eng. **3**, 190–215 (2012).
15. Y. Ling, “The color perception of natural objects: familiarity, constancy and memory.,” in *School of Biology and Psychology*(University of Newcastle, Newcastle upon Tyne: 2005), p. 173.
16. E. Kanematsu and D. H. Brainard, “No measured effect of a familiar contextual object on color constancy,” Color Res. Appl. **39,** 347–359 (2013).
17. A. C. Hurlbert and Y. Ling, “If it's a banana, it must be yellow: The role of memory colors in color constancy,” J. Vis. **5**(8), 787 (2005).
18. T. Hansen, M. Olkkonen, S. Walter, and K. R. Gegenfurtner, “Memory modulates color appearance,” Nat. Neurosci. **9**(11), 1367–1368 (2006).
19. J. S. Bruner, L. Postman, and J. Rodrigues, “Expectation and the perception of color,” Am. J. Psychol. **64**(2), 216–227 (1951).
20. K. Duncker, “The influence of past experience upon perceptual properties,” Am. J. Psychol. **52**(2), 255–265 (1939).
21. C. J. Bartleson, “Color in memory in relation to photographic reproduction,” Photograph. Sci. Eng. **5**, 327–331 (1961).
22. S. N. Yendrikhovskij, F. J. J. Blommaert, and H. Ridder, “Color reproduction and the naturalness constraint,” Color Res. Appl. **24**(1), 52–67 (1999).
23. P. Bodrogi and T. Tarczali, “Color memory for various sky, skin, and plant colors: Effect of the image context,” Color Res. Appl. **26**(4), 278–289 (2001).
24. P. Bodrogi and T. Tarczali, “Investigation of Color Memory,” in *Color Image Science: Exploiting Digital Media*, L. W. MacDonald, and M. R. Luo, eds. (John Wiley & Sons Limited, 2002), pp. 23–48.
25. H. Zeng and R. Luo, “Modeling memory color region for preference color reproduction,” in *SPIE: Color Imaging XV: Displaying, Processing, Hardcopy, and Applications* (2010), pp. 752808–752808–752811.
26. T. Yano and K. Hashimoto, “Preference Index for Japanese Complexion Color under Illumination,” J. Light Vis. Env. **22**, 54 (1998).
27. C. Boust, H. Brettel, F. Viénot, G. Alquié, and S. Berche, “Color enhancement of digital images by experts and preference judgments by observers,” J. Imaging Sci. **50**(1), 1–11 (2006).
28. S. Xue, M. Tan, A. McNamara, J. Dorsey, and H. Rushmeier, “Exploring the use of memory colors for image enhancement,” in *SPIE: Human Vision and Electronic Imaging XIX* (2014), pp. 901411–901411–901410.
29. C. L. Sanders, “Assessment of color rendition under an iIlluminant using color tolerances for natural objects,” J. Illum. Eng. **54**, 640–646 (1959).
30. K. A. G. Smet, W. R. Ryckaert, M. R. Pointer, G. Deconinck, and P. Hanselaer, “A memory color quality metric for white light sources,” Energy Build. **49**, 216–225 (2012).
31. T. Tarczali, D.-S. Park, P. Bodrogi, and C. Y. Kim, “Long-term memory colors of Korean and Hungarian observers,” Color Res. Appl. **31**(3), 176–183 (2006).
32. S. Fernandez, M. D. Fairchild, and K. Braun, “Analysis of observer and cultural variability while generating “preferred” color reproductions of pictorial images,” J. Imaging Sci. Technol. **49**, 96–104 (2005).
33. C. Sik-Lányi, “Styles or cultural background does influence the colors of virtual reality games?” Acta Polytechnica Hungarica **11**, 97–119 (2014).
34. L.-C. Ou, M. Ronnier Luo, P.-L. Sun, N.-C. Hu, H.-S. Chen, S.-S. Guan, A. Woodcock, J. L. Caivano, R. Huertas, A. Treméau, M. Billger, H. Izadan, and K. Richter, “A cross-cultural comparison of color emotion for two-color combinations,” Color Res. Appl. **37**(1), 23–43 (2012).
35. X.-P. Gao, J. H. Xin, T. Sato, A. Hansuebsai, M. Scalzo, K. Kajiwara, S.-S. Guan, J. Valldeperas, M. J. Lis, and M. Billger, “Analysis of cross-cultural color emotion,” Color Res. Appl. **32**(3), 223–229 (2007).
36. M. Saito, “Comparative studies on color preference in Japan and other Asian regions, with special emphasis on the preference for white,” Color Res. Appl. **21**(1), 35–49 (1996).
37. A. Choungourian, “Color preferences and cultural variation,” Percept. Mot. Skills **26**(3), 1203–1206 (1968).
38. S. Shoyama, Y. Tochihara, and J. Kim, “Japanese and Korean ideas about clothing colors for elderly people: Intercountry and intergenerational differences,” Color Res. Appl. **28**(2), 139–150 (2003).
39. CIE16x-2004, “A review of chromatic adaptation transforms,” (CIE, Vienna, 2004).
40. M. Melgosa, P. A. García, L. Gómez-Robledo, R. Shamey, D. Hinks, G. Cui, and M. R. Luo, “Notes on the application of the standardized residual sum of squares index for the assessment of intra- and inter-observer variability in color-difference experiments,” J. Opt. Soc. Am. A **28**(5), 949–953 (2011).
41. H. Wang, G. Cui, M. R. Luo, and H. Xu, “Evaluation of color-difference formulae for different color-difference magnitudes,” Color Res. Appl. **37**(5), 316–325 (2012).
42. P. E. Shrout and J. L. Fleiss, “Intraclass correlations: Uses in assessing rater reliability,” Psychol. Bull. **86**(2), 420–428 (1979).

43. H. Motulsky and A. Christopoulos, *Fitting Models to Biological Data using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting* (Oxford University, 2004).

1. Introduction

The concept of *memory color* refers to the *color associated with a familiar object in (long-term) memory;* or, as Hering stated in the late 19th century, "*the color in which we have most consistently seen the external object"* and which is "*impressed indelibly on our memory*" [1]. It should be distinguished from color memory, which is the ability to recollect colors in general.

Memory colors are however not perfect mental representations of the original object colors. Bartleson [2] and Newhall, Burnham and Clark [3] found that saturation and brightness tended to increase in memory colors and that hue tended to shift towards the dominant hue within the object for some objects. Smet, Ryckaert, Pointer, Deconinck and Hanselaer [4], while investigating color appearance tolerances for familiar objects, reported similar increases in saturation for memory colors and shifts towards the dominant hue for most familiar objects. Vurro, Ling and Hurlbert [5] also found hue shifts in memory colors of natural objects, but these were not systematically towards the dominant hue of the object. They also found that hue shifts were reduced by increasing the naturalness of the stimuli. Siple and Springer [6] confirmed the tendency for saturation increase, but reported quite accurate agreement for brightness and hue. In a study by Pérez-Carpinell, de Fez, Baldoví and Soriano [7] memory saturation only increased for high purity objects, while it decreased or remained the same for midrange or low purity objects. They also reported unsystematic hue shifts specific to the familiar object investigated. Memory color saturation was also higher for the familiar object – a yellow banana – in the study by Yendrikhovskij, Blommaert and de Ridder [8].

Studies on preferred object colors have also reported saturation increases with respect to the actual object colors [6, 9–12]. The preferred color of an object is not necessarily identical to its memory color [13], although Siple and Springer [6] could not identify any significant differences between memory and preferred colors for food objects.

Memory and preferred colors have long been of interest to many different areas of color research. They have been investigated as a possible mechanism to improve the color constancy of other objects, as they provide cues to the visual system to help estimate the illumination. Although, Granzier and Gegenfurtner [14] reported a small improvement, neither Ling [15], nor Kanematsu and Brainard [16] identified such effect. Ling [17], as well as others [18–20] did however report an influence of the memory color on the perceived color of the familiar object itself, consistent with Hering's statement that "*All objects that are already known to us from experience, or that we regard as familiar by their color, we see through the spectacles of memory color." [1]*.

Memory and preferred colors have also been suggested or used as an internal reference to assess object color appearance and color quality in color reproduction [21–28] and color rendering [11, 12, 29, 30].

When using memory colors as a reference, an important question to answer is: "***Are memory colors geographically dependent?***". Although, there are quite a few studies available investigating cross-regional or cross-cultural influences on color perception or color preference [31–38], the number of them directly related to memory colors is rather limited. In fact only one [31] was found in literature, but memory colors for only two regions – Central-Europe (Hungary) and South-East Asia (Korea) – were investigated. Statistically significant differences were found for many of the memory colors.

The current study is an attempt to contribute to the limited literature on the subject of cross-regional variation of memory colors. The memory colors of a set of 11 familiar objects – and the observer response to any color deviation from them – was determined and analyzed for test subjects from seven different countries/regions.

2. Methods

An international collaboration was set up to study cross-regional (or cross-cultural) variation of memory colors. Seven different laboratories located in respectively Belgium, Hungary, Brazil (State of São Paulo), Colombia (Cundinamarca), China (Shanghai), Taiwan and Iran – covering a large portion of the globe – participated in this study. Eleven familiar objects covering the entire hue circle were selected. Each laboratory determined the memory colors, and the response to a deviation therefrom, using an identical experimental setup and procedure. The experiments were performed in each laboratory on a carefully calibrated monitor.

The following sections describe the choice of familiar objects, the observer panels and the experimental setup and procedure in more detail.

2.1 Familiar objects

Eleven familiar objects – illustrated in Fig. 1 – were selected for this study: green apple (GA), ripe banana (RP), ripe lemon (RL), cauliflower (CA), orange (OR), strawberry (SB), tomato (TO), dried lavender (DL), smurf® (SM), Caucasian skin (SC) and Asian skin (AS).

They were specifically chosen to cover the entire hue circle and because of their familiarity across cultures. Although, it should be noted that a 'smurf®' and 'dried lavender' were found to be unfamiliar objects by the test subjects in Iran (the lab in Iran only joined the collaboration at a later stage), hence no Iranian memory colors could be determined for these objects.

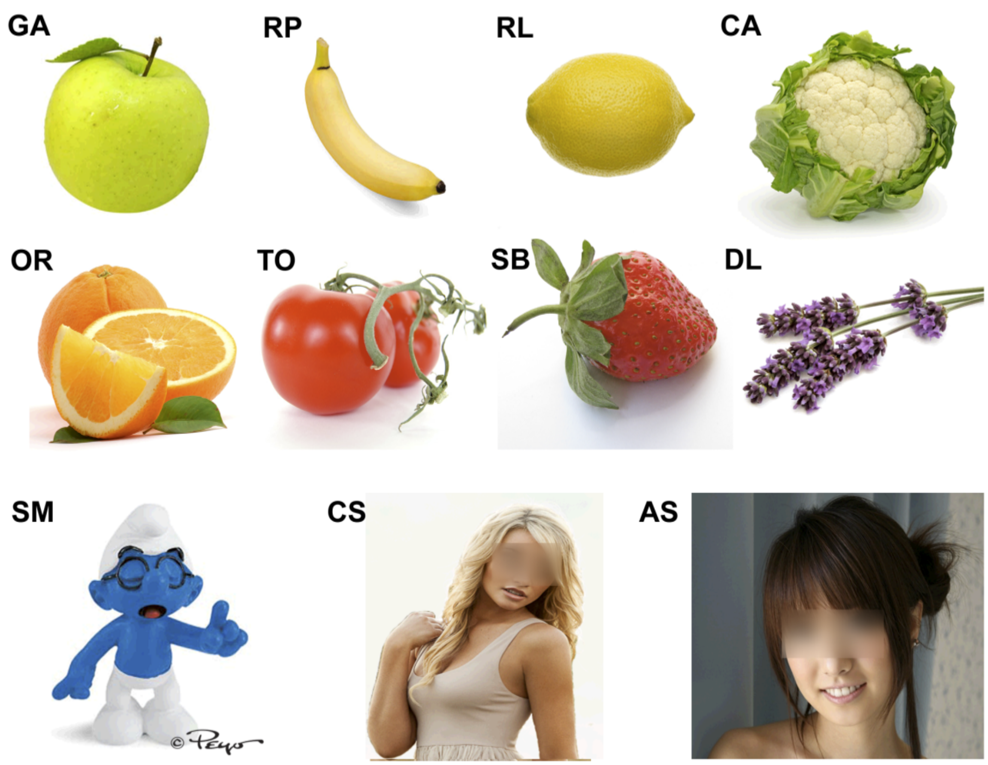

Fig. 1. The eleven familiar 'objects' for which memory colors were determined. Note that images CS and AS are blurred for publication purposes only, no blurring was present during the experiments.

#221779 - $15.00 USD Received 14 Oct 2014; revised 3 Dec 2014; accepted 3 Dec 2014; published 23 Dec 2014

2.2 Test panels

Only color normal test subjects were included in this study. Color deficiency was examined using either Ishihara plates, the Farnsworth-Munsell 100 Hue test or the Farnsworth D15 test following the guidelines specific to each test. In addition to the color normality requirement, test subjects had to be familiar with the object (color) they would be presented with. The goal was to gather visual data from a test panel of 15 observers or more and with an approximate 50/50 male-to-female ratio. Note that the latter varied substantially across laboratories. A total of 280 unique observers participated globally: 28 in Belgium, 14 in Hungary, 20 in Brazil, 99 in Colombia, 19 in Taiwan, 42 in China and 71 in Iran.

The average and minimum number of unique participants per object, the number of total unique observers, the average male-to-female ratio and average age of the test panels for each lab are summarized in Table 1.

Table 1. Summary of the Details of the Test Panels in Each Region

Region	Average # of observers per object	Minimum # of observers per object	Total # of unique observers	Average male-to-female ratio	Average age	Color deficiency test
Belgium	16.2 ± 1.5	15	28	1.24 ± 0.30	31.5 ± 9.6	24 plate Ishihara
Hungary	14.0 ± 0.0	14	14	3.70 ± 0.00	25.7 ± 2.6	24 plate Ishihara
Brazil	15.0 ± 0.5	15	15	0.50 ± 0.00	28.4 ± 5.5	24 plate Ishihara
Colombia	16.0 ± 1.8	14	99	2.81 ± 1.16	24.7 ± 7.1	Farnsworth Munsell 100 Hue
Taiwan	19.0 ± 0.0	19	19	5.33 ± 0.00	22.7 ± 1.01	Farnsworth D-15
China	12.5 ± 1.2	10	42	0.64 ± 0.32	23.2 ± 1.9	24 plate Ishihara
Iran	11.6 ± 1.2	9	71	0.77 ± 0.37	31.9 ± 7.4	38 plate Ishihara

2.3 Experimental procedure

Visual data was collected using an experimental method similar to the one used by Smet, Ryckaert, Pointer, Deconinck and Hanselaer [4]. Test subjects were presented with a familiar object shown in a large number of different colors approximately uniformly spaced in the *CIE 1976 u'v' chromaticity diagram* (calculated using the CIE 1964 observer). They were asked to rate the color appearance of each of the presented object stimuli on a continuous graphical rating scale with respect to what they thought the object looks like in reality, i.e. with their long-term memory color. The scale ran from *very bad* (−1) to *very good* (+ 1) agreement.

Afterwards, the mean (across observers within a single region) color appearance ratings were modeled using a modified bivariate Gaussian function:

$$d\left(u_{10}^{'},v_{10}^{'}\right)=\left\{\left[\begin{pmatrix}u_{10}^{'}\\v_{10}^{'}\end{pmatrix}-\begin{pmatrix}a_3\\a_4\end{pmatrix}\right]^T\begin{bmatrix}a_1 & a_5\\a_5 & a_2\end{bmatrix}\left[\begin{pmatrix}u_{10}^{'}\\v_{10}^{'}\end{pmatrix}-\begin{pmatrix}a_3\\a_4\end{pmatrix}\right]\right\}^{1/2} \tag{1a}$$

$$R\left(u_{10}^{'},v_{10}^{'}\right)=a_7+a_6\cdot e^{-\frac{1}{2}d\left(u_{10}^{'},v_{10}^{'}\right)^2} \tag{1b}$$

with a_{7-7} fitting parameters; $d\left(u_{10}^{'},v_{10}^{'}\right)$ the Mahalanobis distance – which defines an elliptical contour of equal ratings – and $R\left(u_{10}^{'},v_{10}^{'}\right)$ the rating at the chromaticity coordinate $\left(u_{10}^{'},v_{10}^{'}\right)$. The memory color is determined by the centroid $\left(a_3,a_4\right)$ of the bivariate rating function $R\left(u'_{10}, v'_{10}\right)$. The parameters a_1, a_2 and a_5 determine its size, shape and orientation. The $R\left(u'_{10}, v'_{10}\right)$ function provides a convenient description of the observer response to a deviation from the memory color.

The subscript '10' denotes the use of the CIE 1964 (10°) observer in the calculation of the chromaticity coordinates.

2.4 Experimental setup

Smet, Ryckaert, Pointer, Deconinck and Hanselaer [4] performed color appearance rating experiments using real objects presented in a special viewing booth. The color of a familiar object was changed by illuminating it with various settings of RGBY LEDs. The design of the viewing booth masked all cues to the color of the illumination, providing the illusion that the familiar object itself changed color.

However, the construction of such a viewing booth at several locations around the world would be costly and unpractical. Therefore, in this study the familiar objects were presented on a carefully calibrated monitor.

Experiments were performed in a fully darkened room with the monitor as the only source of light. Observers were seated approximately 80 cm from the monitor. Stimuli were presented using a software package especially written for this study. The stimuli were displayed at the center of the monitor and surrounded by a white background to ensure a constant adaptation state to the monitor white point. It is composed of a monitor calibration and stimulus presentation program. The experiment software was sent to each laboratory along with specific instructions to set up the experiment using their own monitor. Care was taken to have the experimental conditions across the different laboratories as identical as possible.

2.4.1 Monitor calibration

The monitor white point was set as close as possible to a D65 chromaticity (u'_{10}, v'_{10} = 0.1979, 0.4695) at a luminance of Y_{10} = 200 cd/m^2. During monitor calibration, a set of RGB stimuli of approximately the same size as the familiar object stimuli was presented in the center of the screen at the location. After spectral measurement and calculation of the XYZ_{10} tristimulus values of the RGB stimuli, the monitor calibration software generated a set of calibration parameters. These parameters included black point, white point, tone response curves and 3 × 3 matrices to go from RGB to *XYZ* and back. The *R*, *G* and *B* tone response curves were obtained with respect to the CIE 1964 *L*, *M* and *S* cone responses, as preliminary tests with several LCD monitors had shown to give better color accuracy than the usual luminance or principal components approach. The stimulus presentation software then used those parameters to present on each monitor, within its calibration accuracy and monitor gamut, the same set of colored stimuli for a familiar object. The accuracy of the calibration was assessed by generating 40 random test colors within the monitor gamut at three distinct uniformly spaced luminance levels and calculating the average and maximum $\boldsymbol{\Delta E^*_{lab}}$ color difference between the target stimuli and the (spectrally) measured stimuli. Some of the details of the calibration, like monitor type, white point setting, luminance, mean and maximum color difference are summarized in Table 2. The monitor gamuts and white points, plotted in the *CIE 1964 chromaticity diagram*, for the monitors used by the laboratories are shown in Fig. 2.

#221779 - $15.00 USD Received 14 Oct 2014; revised 3 Dec 2014; accepted 3 Dec 2014; published 23 Dec 2014

Table 2. Monitor Details: Type, Brand, Gamut, the 10° CIE 1976 *u'v'* Chromaticity Coordinates and Y_{10} Luminance of the Monitor White Point; Mean, Standard Deviation (SD) and Maximum Calibration Error ΔE^*_{Lab}, for a Set of 40 Random Test Chromaticities within the Monitor Gamut

Region	Monitor type, brand, gamut	White point			Calibration error ΔE^*_{lab}	
		u'_{10}	v'_{10}	Y_{10} cd/m^2	Mean ± 1 SD	Max
Belgium	LCD, EIZO, wide gamut	0.1979	0.4698	206	0.96 ± 0.55	2.27
Hungary	LCD, ACER GD245HQ	0.1981	0.4681	201	1.27 ± 0.50	2.39
Brazil	LCD, LG, W1941S	0.1950	0.4724	201	2.65 ± 1.24	5.71
Colombia	LCD, NEC, wide gamut	0.1982	0.4694	200	1.16 ± 0.54	2.35
Taiwan	LCD, Apple, sRGB gamut	0.1980	0.4757	205	2.84 ± 1.33	5.79
China	LCD, EIZO, wide gamut	0.1994	0.4675	227	1.92 ± 0.60	3.29
Iran	LCD, EIZO, wide gamut	0.1982	0.4704	219	1.10 ± 0.48	2.36

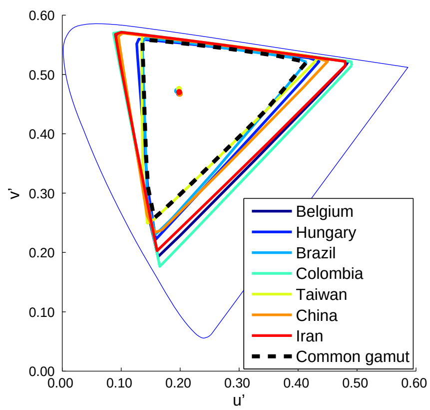

Fig. 2. Monitor gamut and white point in the CIE 1964 chromaticity diagram of the participating laboratories. The common gamut is also plotted.

2.4.2 Stimulus presentation and rating

At each laboratory color appearance ratings were collected for each familiar object using the same stimulus presentation program. Familiar objects were presented in the center of the screen in a large number of different colors approximately uniformly spaced in the *CIE 1976 u'v' chromaticity diagram,* while keeping the luminance nearly constant. Based on preliminary experiments, the extent of grid of test points was chosen such that object colors rated 'acceptable-to-very good' were maximally surrounded by those rated 'very bad', as this minimizes possible bias when fitting the bivariate Gaussian models to the observer ratings [4]. During stimulus presentation, only those areas (pixels) associated with the prototypical color of the familiar objects were changed by using predetermined *template images* that identified the pixels to be altered. The original luminance values of the familiar object images were kept intact and only the chromaticity (of the target pixels) was changed. Two examples of a typical screen as seen by a test subject during the experiment are shown in Fig. 3. The large white area surrounding the familiar object was to ensure adequate and constant adaptation to a D65 chromaticity by minimizing adaptation to the stimulus itself. To further avoid the latter, test subjects were also instructed NOT to stare at the object. The average luminance values of the experimental stimulus grids are given in Table 3. The stimuli

#221779 - $15.00 USD Received 14 Oct 2014; revised 3 Dec 2014; accepted 3 Dec 2014; published 23 Dec 2014

coordinates were also corrected for the slight deviation of the monitor white point from target D65 chromaticity by the CAT02 chromatic adaptation transformation [39].

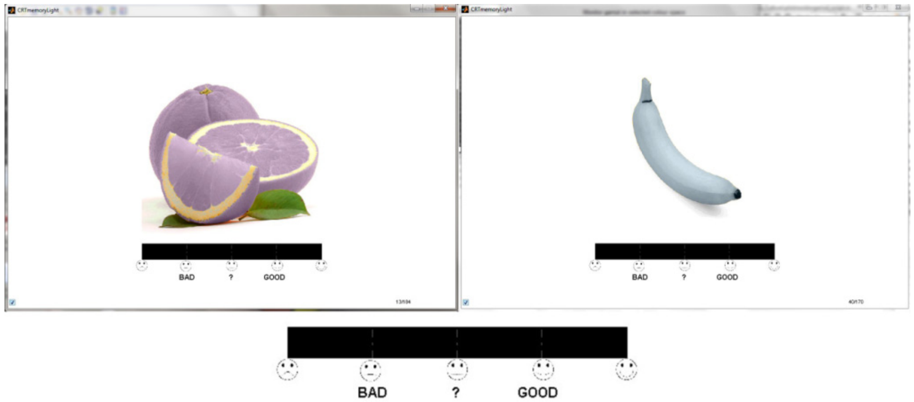

Fig. 3. Top left and top right: two examples of a typical screen as seen by a test subject during the experiment. Bottom: a close up of the continuous graphical rating scale.

Table 3. Mean Luminance Values (Y_{10} in cd/m^2) and the Standard Deviation (SD) for the Different Object Stimulus Grids

Luminance (cd/m^2)	Apple	Banana	Lemon	Cauliflower	Orange	Tomato	Strawberry	Lavender	Smurf ®	Caucasian Skin	Asian skin
Mean	121.2	122.4	87.2	124.0	90.0	67.4	44.1	24.4	25.8	53.6	54.8
SD	6.5	6.8	5.2	6.1	4.5	3.7	2.4	10.9	11.5	2.7	2.8

The exact number of presented stimuli varied from object to object and from laboratory to laboratory, as only stimuli within the laboratory's monitor gamut were selected. They are listed in Table 4. The average number of ratings per object, whereby each observer rated each object color only once per session, was 165 ± 24. Taking into account the number of test subjects that participated, a total of over 210000 ratings have been made during the course of this study, with an average of about 30000 per region.

Table 4. Number of Presented Stimuli per Object for Each Region

Region	**Apple**	**Banana**	**Lemon**	**Cauliflower**	**Orange**	**Tomato**	**Strawberry**	**Lavender**	**Smurf ®**	**Caucasian Skin**	**Asian skin**	**Mean**
Belgium	174	170	154	172	184	208	215	192	189	180	169	**182**
Hungary	146	134	115	140	143	146	164	152	146	180	153	**147**
Brazil	144	129	111	127	127	128	144	146	145	180	145	**139**
Colombia	175	171	158	172	188	216	225	199	198	180	170	**187**
Taiwan	164	151	127	144	149	149	164	140	147	180	164	**153**
China	171	166	149	168	178	185	181	163	162	180	165	**170**
Iran	175	171	155	172	185	207	211	*0*	*0*	180	169	**181**
Mean	**164**	**156**	**138**	**156**	**165**	**177**	**186**	**165**	**165**	**180**	**162**	**165**

#221779 - $15.00 USD Received 14 Oct 2014; revised 3 Dec 2014; accepted 3 Dec 2014; published 23 Dec 2014

For each stimulus the test subject rated the color appearance with respect to how he/she thought the familiar object looks like in reality by clicking on a continuous graphical rating scale presented below the familiar object (see Fig. 3).

The assessment of one object took about 15 - 20 minutes, not including instructions. Test subjects were allowed to rate more than one object a day. However, repeats – to asses intra observer variability – were performed on separate days.

3. Results and discussion

3.1 Observer variability

Intra-observer variability was assessed with STRESS, the Standardized-Residual-Sum-of-Squares [40]. Higher values indicate higher variability (less agreement). The intra-observer variability was obtained by having 2 or more observers repeat a color appearance rating experiment one or more times on separate days. First, for each observer (and object) a STRESS value was calculated between the individual repeated rating sets and their mean. Secondly, a general STRESS value was obtained by averaging the former across observers and objects. Note that not all laboratories had intra-observer data for all objects: BE, BR, TW and CN had 0 missing, CO and HU had respectively 1 and 5 missing, while IR had only one intra-observer data set. The general average intra-observer variability values and their standard deviations are given in Table 5.

Table 5. Average Intra- and Inter-observer Variability[a]

region	Intra-STRESS Mean ± 1SD	Inter-STRESS Mean ± 1SD	ICC(2,k) Mean ± 1SD
BE	0.20 ± 0.03	0.34 ± 0.04	0.95 ± 0.02
HU	0.25 ± 0.05	0.38 ± 0.05	0.96 ± 0.02
BR	0.26 ± 0.06	0.40 ± 0.05	0.93 ± 0.03
CO	0.26 ± 0.03	0.36 ± 0.03	0.93 ± 0.03
TW	0.25 ± 0.04	0.41 ± 0.02	0.89 ± 0.08
CN	0.17 ± 0.01	0.27 ± 0.03	0.96 ± 0.01
IR	0.19[b]	0.35 ± 0.04	0.89 ± 0.07

[a]The standard deviation is also given.

[b]No standard deviation could be calculated for the Iranian data as data for only 1 object was available.

The values ranged from 0.17 to 0.26, with an average of 0.22 ± 0.03. The STRESS values, typically also found in color discrimination studies, indicate a satisfactory agreement between individual observer results obtained on separate days. The degree of intra-observer variability was mostly very similar for all familiar objects tested, as is shown by the generally small standard deviations in Table 5. IR reported data for only one object, a ripe apple. Therefore no standard deviation could be calculated.

Inter-observer variability was also assessed with the STRESS measure. For each object, it was calculated between the mean (across repeats) individual observer ratings and the ratings of the average observer (mean across individual observers). The mean (across objects) STRESS values for the different regions are shown in Table 5. The values range from 0.27 to 0.41, with an average of 0.36 ± 0.04. These STRESS values are typical for inter-observer variation in color difference studies [40, 41], which is remarkably good considering the test subjects did not rate with respect to a single reference stimulus, but to his/her own non-physical memory color. The degree of inter-observer variability was quite similar for the different objects, as indicated by the relatively small standard deviations. The STRESS results also show the inter-observer variability to be typically about 1.7 times larger than the intra-observer variation.

In addition to the STRESS value, inter-observer variability was also evaluated by calculating the *ICC(2,n)* Intraclass Correlation Coefficient [42], which expresses the reliability of the concept of an *average observer* based on the ratings of a limited number of individual observers.

While the good inter-observer variability indicated the validity to calculate an average observer to represent the panel of individual observers, the "excellent" (≥ 0.90) *ICC(2,n)* values – shown in Table 5 – validate the extension to a more general average observer representing the population from which the test panel was randomly drawn. In other words, the average observer obtained by taking the mean of the individual observer ratings should not only be representative of the test panel, but also of the entire population (region/culture).

3.2 Modeling the color appearance ratings

First, for each region and each object an average observer was calculated by taking the mean of the individual observer ratings. An example of the distribution in the CIE *u'v'* chromaticity diagram of the average observer ratings for each region for the familiar object "Asian Skin" is shown in Fig. 4. The distribution corresponding to the set of pooled ratings (of all regions) is also plotted. It's clear that the rating distributions of the regions are quite similar, especially in terms of their overall orientation. The centroids of the distributions, corresponding to the most likely location of the memory color of the object, appeared to be closely grouped (and can be approximated by an elliptical distribution). Even their size is comparable, except for Taiwan.

Following Smet, Ryckaert, Pointer, Deconinck and Hanselaer [4] the mean rating scores were therefore modeled by a bivariate Gaussian function *R(u',v')* (see Eq. (1). The fitting parameters $a_{1\text{-}7}$ for each of the rating functions *R(u',v')* are given in Appendix A. For Iran no rating data for smurf® (SM) and dried lavender (DL) are available as these objects are not familiar to Iranian people. As an example, a 3D plot of the fitted models is made for "Asian Skin " in Fig. 5.

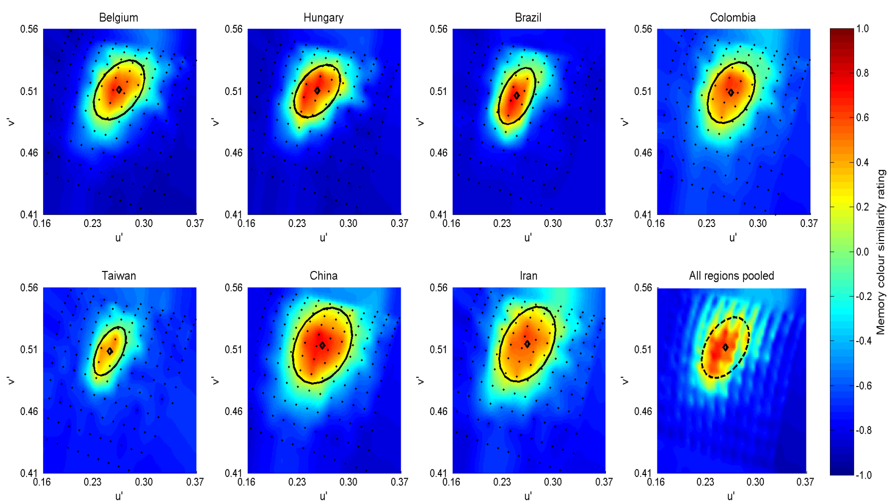

Fig. 4. Average color appearance rating distributions in the CIE 1976 *u'v'* chromaticity diagram for "Asian Skin" for all regions and for all regions pooled. The black dots are the test chromaticity points.

From Fig. 4, it can be observed that the 1*d*-elliptical contours are a good first order approximation for the chromaticity area associated with acceptable – positive – observer ratings. The 1*d*-elliptical contours for all objects and regions are illustrated in Fig. 6.

From Fig. 5, it is clear that there is good agreement between the experimentally determined mean observer ratings and the model fits. The goodness-of-fit was more quantitatively assessed by calculating the STRESS between the mean observer ratings and the modeled ratings. For “Asian Skin” the STRESS values ranged between 0.10 and 0.17, with an average of 0.14 ± 0.03. Compared to the inter-observer variability these values were much smaller, indicating a very satisfactory fit to the rating data. In fact, the agreement of the fit with the average observer is approximately 3 times better than the agreement of a random observer with the average observer!

The other objects had similarly excellent STRESS values as can be seen from the data – shown in percent (%) – in Table 6, where the minimum and maximum STRESS values are 0.08 and 0.17 respectively.

Table 6 also shows the results for the fit to the pooled region rating data. In contrast to the previous results, these STRESS values are much larger. They range from 0.25 to 0.56, with an average of 0.37 ± 0.09. The pooled STRESS values are thus comparable to the ones found for the inter-observer variability within a single region. In other words, the *average observer for a single region* deviates approximately the same amount from the *global average observer* — as an *individual observer within a region* deviates from its average observer.

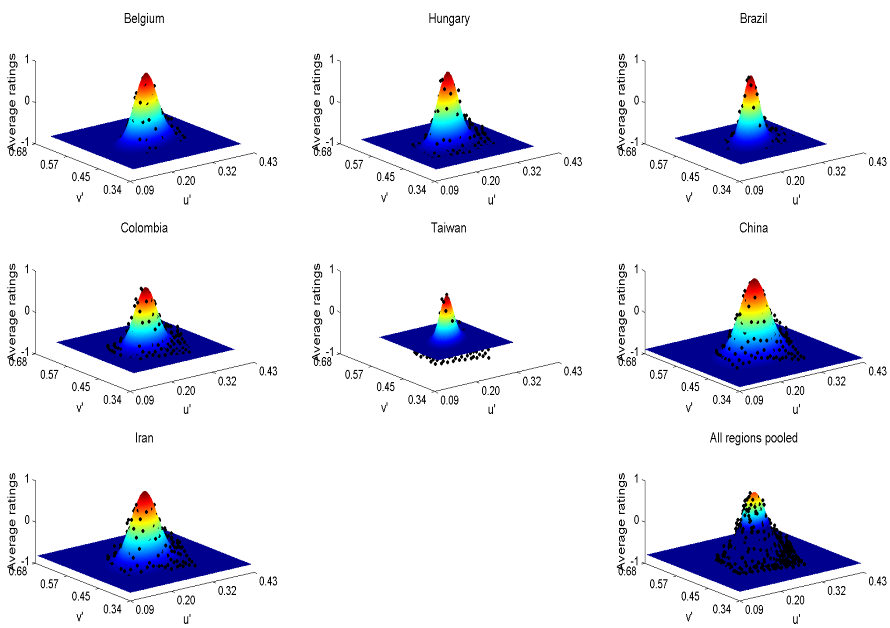

Fig. 5. A 3D plot of the fitted models for “Asian skin” for all regions. The fit to the pooled region data is also shown.

Table 6. Goodness-of-fit of the Bivariate Gaussian Model with the Mean Observer Ratings Expressed in STRESS (%) for the Different Regions and Objects*

object	BE	HU	BR	region CO	TW	CN	IR	Mean	SD	Pooled
GA	12	13	12	16	14	10	13	13	2	56
RB	11	9	12	10	11	8	17	11	3	25
RL	13	12	17	9	13	11	15	13	3	43
CA	15	16	15	13	13	11	13	14	2	35
OR	10	12	12	12	10	10	13	11	1	25
TO	11	12	12	14	12	13	16	13	2	36
SB	10	11	11	16	11	11	11	11	2	31
DL	15	17	14	14	11	14	-	12	6	33
SM	16	17	15	14	13	14	-	13	6	37
CS	12	16	15	16	13	13	14	14	1	47
AS	15	17	17	12	15	10	14	14	3	33

***The mean and its standard deviation (SD) across regions and the STRESS for the fit to the pooled mean ratings (over all regions) are also given. STRESS values are given in percent values.**

#221779 - $15.00 USD Received 14 Oct 2014; revised 3 Dec 2014; accepted 3 Dec 2014; published 23 Dec 2014

3.3 Cross-regional/cultural differences in color appearance rating

Unlike intra-region variability – which is equivalent to the inter-observer variability within a single region – inter-region variability could not be directly evaluated by calculating the STRESS between the mean observer ratings within a single region and the mean observer ratings averaged across all regions, because different regions had slightly different stimuli sets. The sets differed both in size and in chromaticity of the test stimuli. The former due to differences in monitor gamut and the latter due to slightly different white points which resulted in slightly different corresponding chromaticity after correction to a D65 adaptation white. For this reason, cross-regional (or -cultural) differences have been analyzed by comparing – for each object – the region average observers with the global average observer. The former were modeled by the bivariate Gaussians fitted to the mean observer ratings for each region, while the latter was modeled by the bi-Gaussian fitted to the pooled set of mean observer ratings of all regions. Statistical significance of cross-regional effects on color appearance rating and memory colors was evaluated by the extra-sum-of-squares *F*-test [43]. This *F*-test compares the goodness-of-fit of two alternative models, one being a simpler "nested" version of the other. In the analysis at hand, the simple model was the average global observer fit. This model assumes that the variance in the entire rating data set can be explained by a single bivariate Gaussian function with 7 parameters. The other, more complex model, assumes a separate bivariate Gaussian function for each region is required to explain the variance. This model has 49 parameters, 7 for each separate region fit. The null hypothesis is that the simple model is correct. The *F*-test compares the improvement in the residual Sum-of-Squares (*SS*) for the more complicated model with the loss in degrees of freedom (*DF*) associated with the increase in the number of model parameters:

$$F = \frac{\left(SS_{null} - SS_{alt}\right)/\left(DF_{null} - DF_{alt}\right)}{SS_{alt}/DF_{alt}} \tag{2}$$

with SS_{null} and SS_{alt} the residual sum-of-squares between the model fit and the visual data for the simple and complex model respectively. *DF* refers to the degrees of freedom of the simple and complex models.

The results of the extra-sum-of-squares *F*-test (see Table 7) showed a statistically significant effect for all the familiar objects, meaning at least one region average observer differed significantly from the global average observer. Posthoc cross-comparison *F*-tests showed that all regions differed significantly ($p < 0.001$) from one another for all objects, even after Bonferroni correction of the significance level. However, the effect size was small as can be seen from the eta-square value ($\boldsymbol{\eta}^2$) in Table 7.

Table 7. Details of the Extra-sum-of-squares *F*-test for the Effect of Geographical Region

object	F	$F_{\text{crit}}(0.05, df_1, df_2)$	df_1	df_2	p	$\boldsymbol{\eta}^2$
GA	35.6	1.397	42	960	$p < 0.001$	0.19
RB	9.1	1.398	42	903	$p < 0.001$	0.02
RL	33.9	1.400	42	780	$p < 0.001$	0.12
CA	15.4	1.398	42	906	$p < 0.001$	0.05
OR	23.1	1.397	42	965	$p < 0.001$	0.03
TO	28.0	1.396	42	1050	$p < 0.001$	0.07
SB	28.6	1.395	42	1115	$p < 0.001$	0.05
DL	12.0	1.437	35	830	$p < 0.001$	0.04
SM	20.3	1.437	35	825	$p < 0.001$	0.06
CS	33.4	1.396	42	1071	$p < 0.001$	0.12
AS	21.7	1.397	42	946	$p < 0.001$	0.05

As the male-to-female ratio of the test panels varied widely across laboratories, *gender* is a possible confounding factor, i.e. the differences observed could be solely due to differences in rating between men and women. To investigate the gender effect, another extra-sum-of-squares F-test was performed, but this time between the *global average observer* and the *global male* and *female* observers. The latter were obtained by pooling respectively all male and female data from all geographical regions. The F-test results, given in Table 8, show a statistically significant effect for six objects, indicating the global male, global female observer or both did differ significantly from the global average observer for about half of the objects. However, the effect sizes are extremely small compared to those obtained for the effect of geographical region. In fact, on average they were about 30 times smaller. It can therefore be concluded, in agreement with the earlier analysis results, that geographical region did indeed have a significant effect, but that the size of the effect is small.

Table 8. Details of the Extra-sum-of-squares *F*-test for the Effect of Gender

object	F	$F_{crit}(0.05, df_1,df_2)$	df_1	df_2	p	η^2
GA	5.4	2.015	7	1849	$p < 0.001$	0.0097
RB	8.5	2.014	7	1890	$p < 0.001$	0.0061
RL	1.9	2.015	7	1644	$p = 0.070$	0.0028
CA	1.7	2.014	7	1896	$p = 0.106$	0.0017
OR	3.2	2.014	7	2014	$p = 0.002$	0.0015
TO	2.0	2.014	7	2184	$p = 0.057$	0.0014
SB	3.8	2.014	7	2314	$p < 0.001$	0.0026
DL	1.3	2.015	7	1730	$p = 0.228$	0.0011
SM	3.5	2.015	7	1720	$p = 0.001$	0.0044
CS	3.6	2.014	7	2066	$p < 0.001$	0.0052
AS	1.2	2.015	7	1826	$p = 0.306$	0.0010

That the impact of region or culture on color appearance rating and memory color, although significant, is small is also clear from Fig. 6 where the $1d$-elliptical contours of the fitted bivariate Gaussian functions show very similar location, size and orientation.

In fact, as already mentioned during the discussion on the goodness-of-fit of the bivariate Gaussian models and as can be seen from a comparison of Figs. 6 and 7, the observed variability between different regions is of the same order of magnitude or smaller than the typical inter-observer variability within a single region.

Statistically significant, but small – smaller than the intra-culture variation – and hence practically irrelevant effects of culture were also reported by Fernandez, Fairchild and Braun [32] in a study on observer and cultural variability of preferred colors of pictorial images.

#221779 - $15.00 USD Received 14 Oct 2014; revised 3 Dec 2014; accepted 3 Dec 2014; published 23 Dec 2014

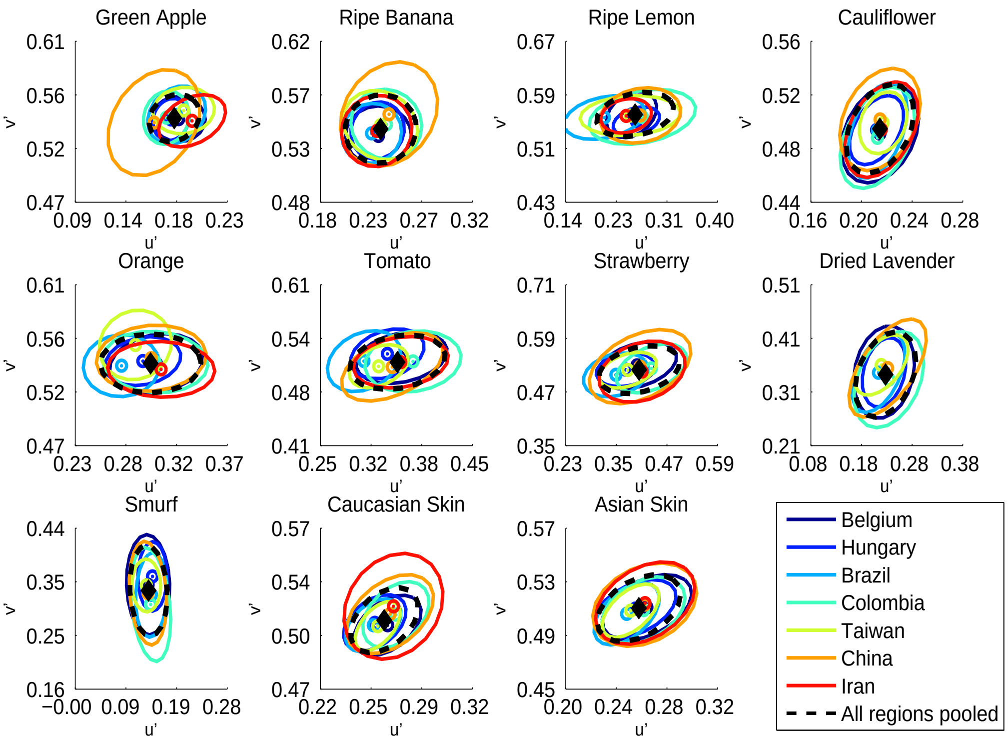

Fig. 6. The 1d-elliptical contours – at a unit Mahalanobis distance – of the fitted bivariate Gaussian models (for the average observer) for each of the 7 regions and 11 objects. The 1d-elliptical contour of model fitted to the pooled data (across regions) is plotted as a dashed black line.

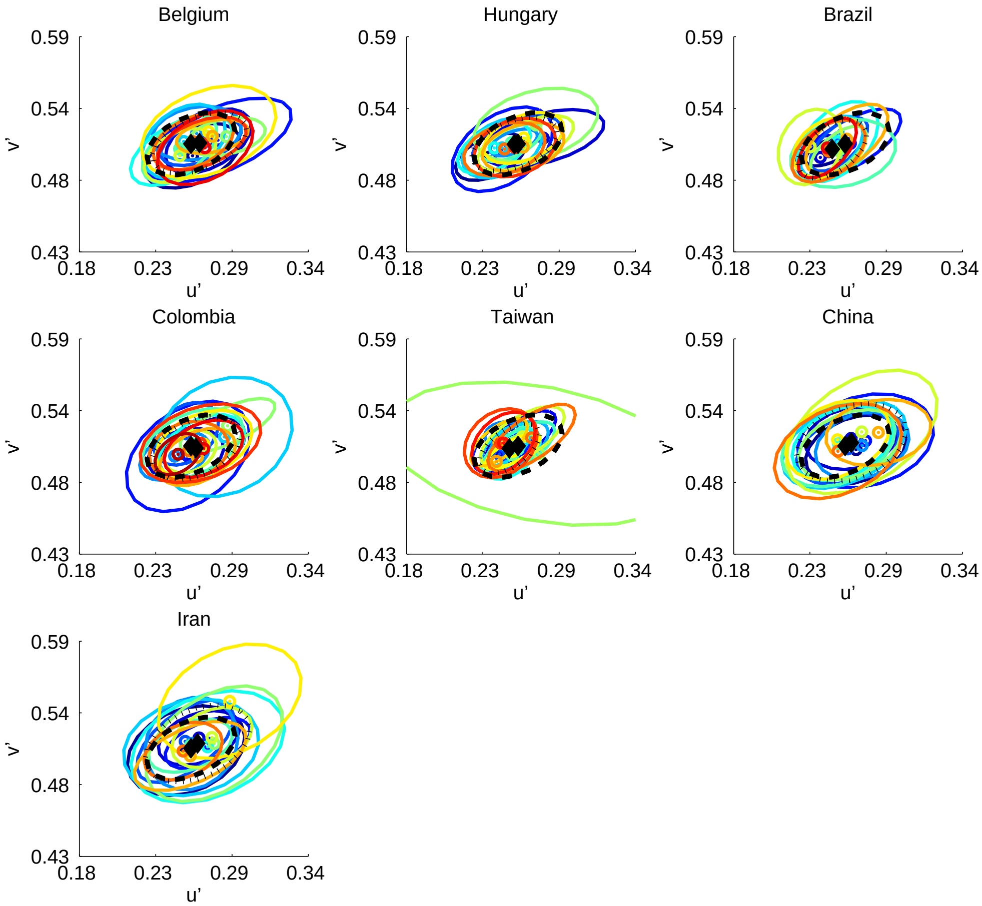

Fig. 7. The 1*d*-elliptical contours of the bivariate Gaussian models fitted to the individual observer data for '*Asian Skin*' (colored solid lines), fitted to the average ratings (across observers) of each region (dotted black line) and fitted to the pooled (across regions) ratings (dashed black line).

#221779 - $15.00 USD Received 14 Oct 2014; revised 3 Dec 2014; accepted 3 Dec 2014; published 23 Dec 2014

Finally, to illustrate the size of the regional effect on memory color in a practical way, the region-average memory colors are displayed for "*Asian Skin*" in Fig. 8 and for "*Green Apple*" in Fig. 9 as an example. The memory color for the global observer – obtained from the pooled region-ratings – is also displayed. In addition, the chromaticities of the displayed memory colors are also shown. The displayed images were calculated using the *srgb* color space. Although, the actual displayed colors will depend on the medium on which they are presented, *srgb* does give a good first approximation on a typical monitor. For "*Asian Skin*" (Fig. 8), it is clear that the differences between the displayed memory colors, although visible, are small indeed.

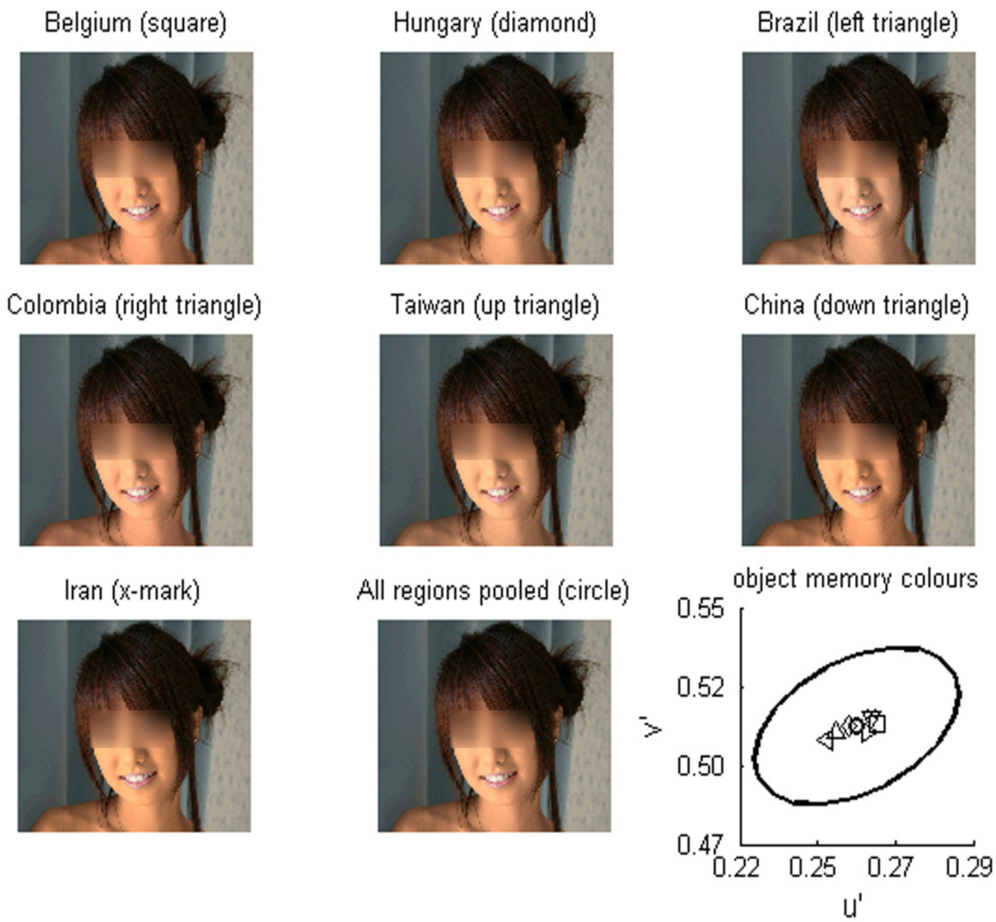

Fig. 8. Region average memory colors for "*Asian Skin*". The memory color for the global observer (all regions pooled) is also shown. The last subplot displays the CIE *u'v'* chromaticity coordinates of the different displayed memory colors and the 1*d*-elliptical contour of the global observer. The display colors were calculated using *srgb* color space. Note that the images are blurred only for publication purposes, no blurring was present during the experiments.

For "*Green Apple*" (Fig. 9), the cross-regional differences in memory colors were more striking. A possible explanation might be the object's large natural variation in color, due to among others, different types of green apples and varying stages of ripeness. Note that all region-average memory colors are still located within the 1*d*-elliptical contour – an approximate tolerance boundary for color acceptability – of the global average observer (see last graph of Fig. 9). Similar graphs for the region-average memory colors of the other familiar objects are plotted in Fig. 10.

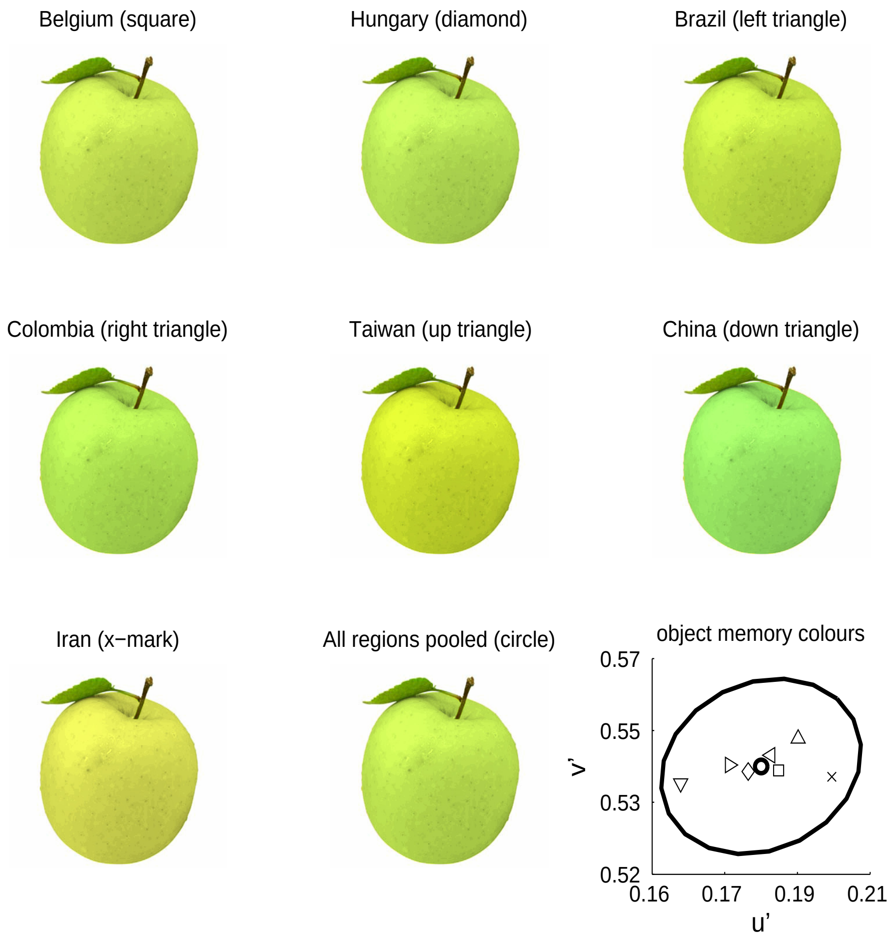

Fig. 9. Region average memory colors for a "*Green Apple*". The memory color for the global observer (all regions pooled) is also shown. The last subplot displays the CIE *u'v'* chromaticity coordinates of the different displayed memory colors and the 1*d*-elliptical contour of the global observer. The display colors were calculated using srgb color space.

#221779 - $15.00 USD Received 14 Oct 2014; revised 3 Dec 2014; accepted 3 Dec 2014; published 23 Dec 2014

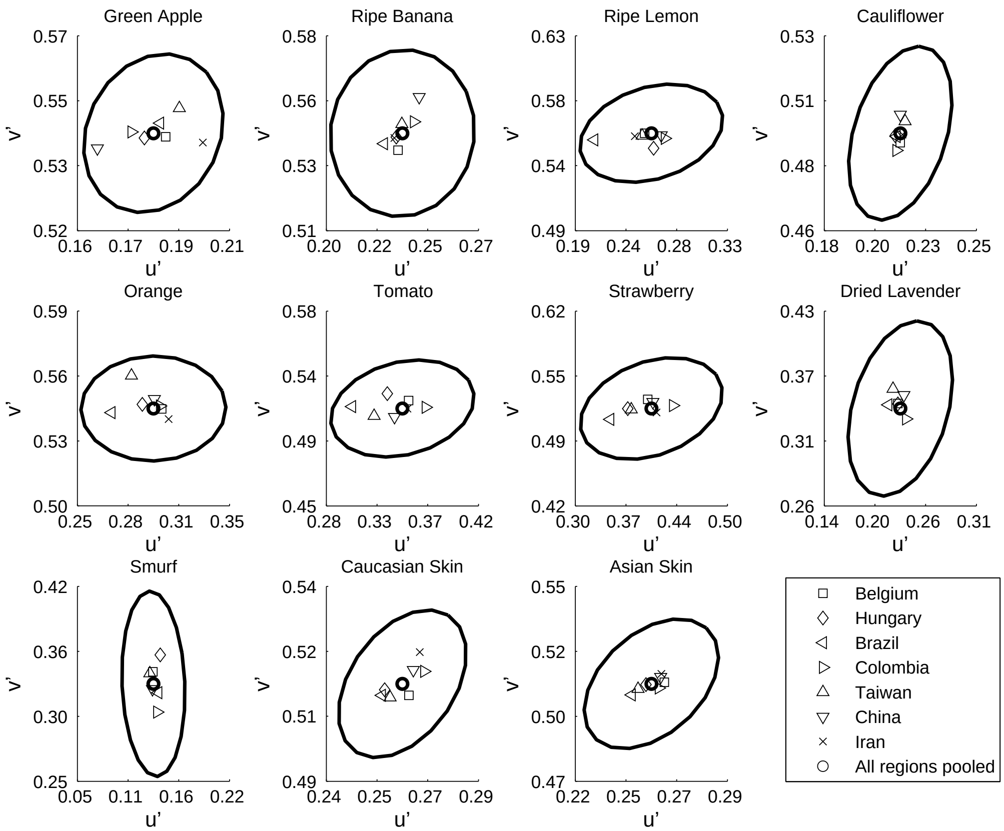

Fig. 10. Region-average memory color chromaticities. The 1*d*-elliptical contour and memory color of the global average observer is also plotted.

The color differences in the CIE 1976 *u'v'* chromaticity diagram between the region average memory colors and the global average memory colors are given in Table 9. The color differences were calculated using the *normalized Mahalanobis distances* – which reduce to a regular Euclidean distance when the *normalized* a_5 parameters are equal to zero – to the global average memory color. By taking the shape and orientation of the global average observer rating function into account, the *normalized Mahalanobis color difference* more accurately represents the actual perceptual difference with the global average memory color than a regular Euclidean distance. The distance in *u'v'* corresponding to one Mahalanobis unit of the global average rating function is also given for comparison. It's clear that the mean color difference is substantially smaller than the unit Mahalanobis distance, indicating, as noted earlier, that all region average memory colors fall well within the acceptability region of the global average observer. It may also be noted that larger variability with respect to the global average (as assessed by the STRESS of the fit with the pooled data in Table 6 or by the $\boldsymbol{\eta}^2$ in Table 7) does not necessarily correspond to larger perceptual differences, as variability and perceptibility are respectively expressed in relative and absolute measures. For example, consider the "*green apple*" and the "*strawberry*". While the variability for the green apple is larger than for the strawberry (see Table 6 and 7), it is also clear from Table 9 that the perceptibility for the green apple is substantially smaller, due to the much smaller unit Mahalanobis distance (MD in Table 9). In addition, variability is assessed with respect to observer ratings for any chromaticity (cfr. Gaussian models), while perceptibility is evaluated only with respect to memory color chromaticity. It may also be noted that, overall, the differences in memory color chromaticity between the region-average observers and the global average observers are larger than (or approaching) the perceptual discrimination limit (1 JND ≈$\Delta E_{u'v'}$ ≈0.003). However, as indicated by when discussing the observer and region variability, similar or larger differences also occur between individual observers and the region average observer.

Table 9. Color Differences – Calculated Using the Mahalanobis Distance Metric – in the CIE 1976 *u'v'* Chromaticity Diagram between the Region Average Memory Colors and the Global Average Memory Colors*

				regions						
	BE	**HU**	**BR**	**CO**	**TW**	**CN**	**IR**	**µ**	**SD**	**MD**
GA	0.0045	0.0032	0.0033	0.0078	0.0104	0.0193	0.0179	0.009	0.007	0.030
RB	0.0065	0.0031	0.0093	0.0067	0.0034	0.0151	0.0040	0.007	0.004	0.045
RL	0.0059	0.0112	0.0498	0.0134	0.0079	0.0093	0.0144	0.016	0.015	0.064
CA	0.0037	0.0024	0.0018	0.0063	0.0044	0.0072	0.0014	0.004	0.002	0.035
OR	0.0051	0.0071	0.0255	0.0041	0.0207	0.0045	0.0106	0.011	0.009	0.050
TO	0.0068	0.0183	0.0442	0.0201	0.0241	0.0079	0.0044	0.018	0.014	0.064
SB	0.0118	0.0311	0.0520	0.0278	0.0262	0.0064	0.0091	0.024	0.016	0.094
DL	0.0060	0.0057	0.0152	0.0132	0.0219	0.0113		0.012	0.006	0.083
SM	0.0103	0.0268	0.0089	0.0243	0.0095	0.0049		0.014	0.009	0.090
CS	0.0042	0.0052	0.0060	0.0064	0.0040	0.0038	0.0076	0.005	0.001	0.022
AS	0.0061	0.0025	0.0087	0.0045	0.0059	0.0042	0.0048	0.005	0.002	0.033

***The mean (µ) and standard deviation (SD) across the different regions is also shown, as well as the *u'v'* distance corresponding to the Mahalanobis distance unit (MD).**

4. Conclusions

The effect of cross-regional or cross-cultural differences on color appearance ratings and memory colors of eleven familiar objects was investigated in seven different regions – Belgium, Hungary, Brazil, Colombia, Taiwan, China and Iran. In each of the corresponding laboratories, the familiar objects were presented on a calibrated monitor in over 100 different colors to a test panel of observers that were asked to rate the similarity of the presented object color with respect to what they thought the object looks like in reality. For each object and region the mean observer ratings were modeled by a bivariate Gaussian function. The goodness-of-fit, as evaluated by the Standardized-Residual-Sum-of-Squares (STRESS) was much smaller than the inter-observer variability STRESS value. Remarkably, considering the rather virtual and subjective – each observer has his/her own –nature of the reference memory color, the inter-observer variability was comparable to that found in color difference studies that employ a fixed reference. A statistical analysis showed significant ($p < 0.001$) differences between the region average observers and the global average observer obtained by pooling the data from all regions. However, the effect size of region or culture was found to be small. In fact, the differences between the region average observers and the global average observer was found to be of the same magnitude or even smaller than the typical inter-observer variability within one region. Thus, although statistical differences in color appearance ratings and memory between regions were found, they are not likely to be of practical importance.

Appendix A: Model parameters of the bivariate Gaussian functions fitted to the rating data (Table 10)

Table 10. Model parameters for the different objects and countries.

Green Apple	a_1	a_2	a_3	a_4	a_5	a_6	a_7
BE	1516.7	2353.7	0.1865	0.5423	11.5	1.373	−0.687
HU	3290.5	2757.4	0.1791	0.5420	−404.4	1.543	−0.772
BR	1338.0	1480.9	0.1846	0.5461	−310.6	1.299	−0.650
CO	2044.7	1696.3	0.1746	0.5437	−169.5	1.192	−0.596
TW	1134.0	2217.5	0.1913	0.5502	−113.8	0.581	−0.290
CN	546.8	463.5	0.1627	0.5392	−144.8	1.467	−0.733
IR	1092.7	2014.5	0.1995	0.5408	−528.7	0.923	−0.461
Pooled	1730.7	2203.2	0.1823	0.5433	−331.0	1.377	−0.689
Ripe Banana	a_1	a_2	a_3	a_4	a_5	a_6	a_7
BE	1032.5	1396.1	0.2331	0.5373	−6.3	1.636	−0.818
HU	1576.5	1545.8	0.2323	0.5421	−122.2	2.119	−1.060
BR	1244.6	1407.9	0.2265	0.5397	−64.0	2.056	−1.028
CO	897.0	1027.4	0.2405	0.5477	49.2	1.663	−0.831
TW	822.6	1060.0	0.2349	0.5468	130.8	1.476	−0.738
CN	494.9	480.9	0.2430	0.5567	−102.5	1.871	−0.936
IR	902.7	993.0	0.2314	0.5417	−47.7	1.960	−0.980
Pooled	876.9	1077.1	0.2351	0.5435	−40.0	1.846	−0.923
Ripe Lemon	a_1	a_2	a_3	a_4	a_5	a_6	a_7
BE	631.1	869.5	0.2533	0.5610	−224.2	1.444	−0.722
HU	753.2	2290.3	0.2612	0.5506	−429.1	1.717	−0.859
BR	244.2	965.3	0.2083	0.5564	−144.4	1.381	−0.690
CO	138.6	602.0	0.2712	0.5575	−82.1	1.235	−0.617
TW	147.0	1068.8	0.2510	0.5594	−54.3	0.972	−0.486
CN	229.0	559.6	0.2680	0.5597	−60.0	2.052	−1.026
IR	582.2	1487.2	0.2444	0.5590	−171.6	1.206	−0.603
Pooled	265.2	920.6	0.2592	0.5611	−133.9	1.521	−0.760
Cauliflower	a_1	a_2	a_3	a_4	a_5	a_6	a_7
BE	1277.3	790.8	0.2148	0.4908	−316.1	1.649	−0.824
HU	2889.3	1772.1	0.2123	0.4936	−979.1	1.837	−0.918
BR	2413.6	930.1	0.2130	0.4933	−648.1	1.767	−0.884
CO	1476.3	798.7	0.2128	0.4878	−435.3	1.539	−0.769
TW	3224.4	1914.5	0.2172	0.4992	−609.0	1.141	−0.571
CN	1182.1	732.6	0.2149	0.5017	−294.8	1.729	−0.865
IR	1514.4	895.5	0.2159	0.4940	−463.1	1.331	−0.666
Pooled	1740.5	1008.9	0.2148	0.4945	−488.9	1.621	−0.810
Orange	a_1	a_2	a_3	a_4	a_5	a_6	a_7
BE	614.5	1617.8	0.3045	0.5413	−186.5	1.903	−0.951
HU	944.2	2425.5	0.2927	0.5435	−300.8	1.728	−0.864
BR	891.5	1447.2	0.2740	0.5395	−32.2	1.977	−0.989
CO	471.7	1490.9	0.3034	0.5427	67.1	1.883	−0.942
TW	1004.5	1150.7	0.2861	0.5571	−93.0	1.330	−0.665
CN	436.9	1275.2	0.3000	0.5459	−72.7	1.969	−0.985
IR	455.5	1791.2	0.3087	0.5364	44.7	2.029	−1.015
Pooled	515.7	1605.4	0.2995	0.5415	−15.4	1.897	−0.949
Tomato	a_1	a_2	a_3	a_4	a_5	a_6	a_7
BE	300.2	1095.7	0.3568	0.5191	−152.4	1.538	−0.769
HU	469.4	1188.2	0.3382	0.5238	−221.9	1.770	−0.885
BR	463.9	782.4	0.3075	0.5149	−162.7	1.994	−0.997
CO	255.4	738.6	0.3717	0.5145	−61.1	1.625	−0.812
TW	771.8	1667.8	0.3267	0.5086	−403.4	1.101	−0.551
CN	286.9	738.3	0.3446	0.5079	−214.9	1.914	−0.957
IR	267.4	1064.7	0.3558	0.5135	−106.2	1.378	−0.689
Pooled	268.5	960.8	0.3514	0.5136	−115.0	1.683	−0.841

Strawberry	a_1	a_2	a_3	a_4	a_5	a_6	a_7
BE	135.2	429.2	0.3995	0.5295	−86.0	1.596	−0.798
HU	222.1	781.6	0.3731	0.5199	−89.5	2.047	−1.024
BR	187.5	455.7	0.3500	0.5087	−73.7	1.885	−0.943
CO	178.8	414.2	0.4328	0.5230	−111.8	1.544	−0.772
TW	262.5	764.9	0.3778	0.5191	−173.1	1.229	−0.615
CN	85.9	181.4	0.4063	0.5267	−51.5	1.880	−0.940
IR	106.7	239.7	0.4112	0.5158	−43.9	1.846	−0.923
Pooled	124.2	396.6	0.4042	0.5200	−66.5	1.763	−0.882
Dried Lavender	a_1	a_2	a_3	a_4	a_5	a_6	a_7
BE	286.6	141.5	0.2247	0.3471	−37.2	1.531	−0.765
HU	575.9	261.0	0.2246	0.3467	−105.5	1.827	−0.914
BR	471.2	241.9	0.2137	0.3459	−163.2	1.774	−0.887
CO	238.4	143.6	0.2342	0.3333	−64.6	1.592	−0.796
TW	476.9	430.9	0.2192	0.3603	−230.0	1.168	−0.584
CN	265.3	190.2	0.2317	0.3546	−135.2	1.880	−0.940
IR							
Pooled	311.8	182.0	0.2275	0.3426	−79.5	1.684	−0.842
Smurf®	a_1	a_2	a_3	a_4	a_5	a_6	a_7
BE	616.2	128.5	0.1342	0.3421	16.4	1.614	−0.807
HU	1337.1	268.7	0.1425	0.3567	−6.1	1.781	−0.891
BR	1166.7	181.9	0.1403	0.3241	−49.5	1.719	−0.860
CO	769.7	106.6	0.1389	0.3074	77.1	1.816	−0.908
TW	1402.4	463.0	0.1310	0.3408	26.6	1.041	−0.520
CN	847.9	123.6	0.1339	0.3271	58.2	1.983	−0.992
IR							
Pooled	830.3	158.3	0.1349	0.3318	37.0	1.748	−0.874
Caucasian Skin	a_1	a_2	a_3	a_4	a_5	a_6	a_7
BE	3101.9	4331.2	0.2648	0.5106	−1646.6	1.526	−0.763
HU	3438.3	4765.2	0.2571	0.5119	−1677.0	1.674	−0.837
BR	4445.3	5465.5	0.2562	0.5106	−2136.8	1.452	−0.726
CO	2986.4	3363.3	0.2697	0.5165	−1221.4	1.349	−0.674
TW	7717.5	7217.7	0.2588	0.5101	−4013.9	0.818	−0.409
CN	1774.0	2408.9	0.2662	0.5169	−854.2	1.513	−0.756
IR	1184.5	1162.7	0.2683	0.5213	−240.2	1.178	−0.589
Pooled	2897.3	3555.5	0.2627	0.5134	−1302.2	1.472	−0.736
Asian Skin	a_1	a_2	a_3	a_4	a_5	a_6	a_7
BE	999.4	2110.8	0.2642	0.5109	−629.3	1.628	−0.814
HU	1209.9	2684.0	0.2552	0.5100	−777.8	1.730	−0.865
BR	2011.4	2498.9	0.2483	0.5061	−1106.8	1.588	−0.794
CO	1092.8	1894.7	0.2615	0.5087	−533.0	1.412	−0.706
TW	2837.2	3628.2	0.2514	0.5084	−1745.7	1.075	−0.538
CN	661.8	1160.0	0.2624	0.5129	−264.3	1.810	−0.905
IR	767.5	1249.0	0.2627	0.5141	−364.0	1.667	−0.833
Pooled	1127.2	1981.2	0.2578	0.5103	−618.2	1.617	−0.809

Acknowledgments

Author KS would like to thank the Research Foundation Flanders (FWO) for supporting this study through a Postdoctoral Fellowship (12B4913N). BVN recognizes FAPESP (09/54292-7) and CNPq (162576/2013-7) for PhD fellowships. GLDC is grateful for a FAPESP (13/03553-0) PhD fellowship. The authors would further like to thank B.Eng. Juan Felipe Gutierrez for his participation in this work at the Color and Vision Laboratory of the Universidad Nacional de Colombia.