

Re-locating colors in the OSA space

Giulia Paggetti · Guido Bartoli · Gloria Menegaz

Published online: 25 November 2010
© Psychonomic Society, Inc. 2010

Abstract A monolexic color-naming experiment in the Italian language was performed on a sample of 1,014 colors obtained extending the OSA-UCS color system (E-OSA). Colors were rendered on a Mitsubishi Diamond pro 2070 monitor as patches subtending 10° of visual angle in a completely dark room. The 11 universal categories proposed by Berlin and Kay (*Basic color terms: Their universality and evolution*, 1969) were considered. Consensus, consistency and focal colors were determined for each category. Comparisons were performed among the results obtained using Boynton & Olson (B&O), the extended OSA (E-OSA), a subset of it covering almost all the B&O sample (R-OSA) and the Sturges & Whitfield (S&W) sample sets. A good overlap could be observed among the locations of the consensus colors in the {L, j, g} color model between B&O and R-OSA, as well as a close proximity among the centroids of homologue regions for the majority of the classes. The analysis of the location of focal colors versus the centroids suggests that an additional basic term could be missing for the Italian language in the blue category.

Keywords Color naming · OSA

G. Paggetti (✉) · G. Menegaz (✉)
Department of Computer Science, University of Verona,
Strada le Grazie 15,
37134, Verona, Italy
e-mail: giulia.paggetti@univr.it
e-mail: gloria.menegaz@univr.it

G. Bartoli
Department of Information Engineering, University of Siena,
Via Roma 56,
53100, Siena, Italy

Introduction

In their pioneering study in 1969, Berlin and Kay (1969) identified 11 basic colors (*red, green, blue, yellow, orange, purple, pink, brown, gray, white, black*). According to their definition, color terms are operationally defined as *basic* only if *monolexic* and *psychologically salient* for all speakers, but not if restricted in application to narrow classes of objects or included in the signification of other color terms. A few years later, in 1987, Boynton and Olson (1987) were able to locate the color categories in the color space developed by the Optical Society of America's Committee on Uniform Color Scales, the Universal Color System (OSA-UCS) (Man & MacAdam, 1989) by evaluating the {L, j, g} coordinates of the centroids and focal colors of each color class. Their monolexic color-naming experiment also enabled the assessment of the psychological salience of *basic* versus *non-basic* color terms.

Existing research in color naming and categorization primarily reflects two opposing views: the cultural *relativist* view that posits that color perception is greatly shaped by culture specific language associations and learning, and the *universalist* view that emphasizes panhuman shared color processing as the basis for color-naming similarities *within* and *across* cultures.

More specifically, cultural relativism suggests that color categorization and naming is due largely to learned language associations specific to a given culture (Davidoff, Davies, & Roberson, 1999; Kay & Kempton, 1984; Roberson, 2005; Roberson, Davies, & Davidoff, 2000; Saunders & Van Brakel, 1997). Universalism, on the other hand, emphasizes panhuman uniformity in the perceptual processing of color as the basis for color-naming coherence within and across cultures (Hardin, 2005; Kay & Regier, 2003; Kay, Berlin, Maffi, & Merrifield, 1997; Regier, Kay,

& Khetarpal, 2007). The universalist theory states that color categories are organized around universal focal colors corresponding to the best examples, or *prototypes*, of the corresponding categories. According to this assumption, the boundaries of color categories are projected from these universal foci and therefore tend to lie in similar positions in color space across languages. In contrast, the opposing relativist theory view denies that foci are a universal basis for color naming and instead maintains that color categories are defined at their boundaries by local linguistic conventions, which are free to vary considerably across languages.

Some researchers have argued that blends of these two different perspectives would be most appropriate for modeling color-naming phenomena (Dedrick, 1997; Jameson, 2005; Jameson & Alvarado, 2003; Paramei, 2005).

In 2007, Regier et al. (2007) proposed that color naming across languages reflects optimal or near-optimal divisions of an irregularly shaped perceptual color space. In 1997, Jameson and D'Andrade (1997) suggested that, given the irregular shape of the color space, where hue interacts with lightness and saturation, general principles of categorization may account for universals of color naming. The hypothesis was that optimal or near-optimal partitions correspond to observed universals in color naming. This proposal may also accommodate the finding that similar languages sometimes have different boundary placements: such languages may have distinct color-naming systems that differ minimally, if at all, in optimality.

Even though it focuses on the six Hering primaries, this work is particularly relevant since it is an attempt to partition the color space according to the data and based on an objective metric. Regier and Kay proposed modeling the formation of categories by the maximization of an objective function which is representative of the *well formedness* of the categories. This accounts for the similarity of samples belonging to the same category and the dissimilarity of those belonging to different categories.

Other interesting contributions are those of Lin, Lou, MacDonald, and Tarrant (2001a, 2001b, 2001c). Two different experiments were performed. The first was based on an unconstrained color-naming task aiming at investigating color naming in English and Mandarin languages. In the second experiment, a constrained method was employed to map a focal colour or a color volume corresponding to each of the important *basic*, *modifier* and *secondary* terms found in the first experiment. Because it was concerned with name-to-color mapping, a constrained method was used. Of note, the semantics of the term *focal color* was defined as the property of being the *most typical* physical color to represent a color term. Interestingly, in this experiments colors were assigned to labels and not vice versa. Finally, a color-naming model was developed to categorize volumes for each of the 11 basic names in CIELAB color space.

Among the main findings was the observation that, because people tend to use the basic level of colors (focal colors) as cognitive reference points, British and Chinese focal colors should be similar, but the range covered by each color might not be the same between individuals or between cultures.

This is very relevant to our work because we take a similar perspective. In our view, while there is an evidence that focal colors tend to cluster in specific regions of the color space, as predicted by the universalist hypothesis, the between-cultural variability could be accounted for by allowing a suitable level of fuzziness in the definition of the category boundaries, which is in agreement with the relativistic hypothesis.

In a previous work (Menegaz, Le Troter, Sequeira, & Boi, 2007), a color-naming model was developed partitioning the CIELAB space in three-dimensional regions each corresponding to a basic category. Data analysis revealed that the membership function for each category peaked in the center of the region and smoothly degraded towards the borders. This is in agreement with the observations of many researchers and is one of the most consolidated assumptions used in the state-of-the-art color-naming models (Benavente, Vanrell, & Baldrich, 2008; Bleys, 2004; Lammens, 1994). The inherent variability in the position of the category boundaries on a single subject basis clearly emerged from the data and was correctly reproduced by the model.

In this work, we do not address the issue of whether foci and boundaries are stable intra-culturally in general. We focus on the Italian language with the objective of establishing to what extent a color-naming experiment in this language and using a calibrated CRT monitor for color display would be agreement with Boynton and Olsons's results. Due to the lack of previous results for the Italian language, it was not possible to split the cultural difference from the other variables intervening in the two experiments. However, under the assumption mentioned above, the inter-cultural variability can be accounted for by enabling the variability in the category boundaries. Accordingly, results are analyzed in terms of focal colors, category centroids, and consistency and consensus rates.

This paper is organized as follows. The first part summarizes the scientific background in the field. The second part illustrates the rationale of the proposed approach and outlines the differences among the proposed and the original set-ups. The third part describes Experiments 1 (color naming) and 2 (identification of focal colors). Results are discussed in the fourth part, and the last part derives the conclusions.

Background

In this section, the two pillar experiments that inspired the proposed work are briefly revisited, and the differences and

similarities of the proposed work with respect to these seminal ones taken as references are discussed.

Boynton and Olson's (B&O) experiment

Boynton and Olson's monolexemic color-naming experiment was designed with the twofold goal of (1) investigating the hypothesis of the existence of basic and non basic color terms and (2) defining the locations of the 11 basic surface colors within the OSA-UCS space.

The positions of the focal colors and the boundaries of each of the 11 regions denoted by such terms in the OSA color system (Boynton & Olson, 1987) were quantitatively determined and the results corroborated Berlin and Kay's hypothesis. The task consisted in naming colors unrestrainedly using a single-word (monolexemic) name.

Seven subjects participated in the experiment. A total of 424 colors were presented one at a time, twice each, for a total of nearly 6,000 observations. Colors were presented in a random order that differed for each observer. All the subjects except one (subject CXO) were instructed to name each color using any monolexemic color term.

From the subjects' responses, the following information was extracted: (1) consistency: agreement on color naming by a single subject in two presentations; (2) consensus: agreement of all subjects in naming a color with a single term in all observations (i.e. consensus implies consistency); (3) position of focal colors, defined as those samples *named with consensus*, that exhibited the shortest response times within their categories; and (4) position of centroids, defined as the centers of the categories and indices of central tendency for color name usage. B&O argued that the response time is a significant index of color saliency, and can be used for distinguishing "basic" terms, named with consensus, from "other basic terms", lacking consensus but named consistently, and "non-basic" terms for which consensus was not reached and consistency was not guaranteed.

The reported consensus, foci and centroids will be compared to those that were obtained in the proposed experiments in the results part.

Sturges and Whitfield's (S&W) experiment

In this experiment, the rendering method was the same as that followed by B&O: 446 Munsell samples were viewed within an enclosure under an illuminant approximating CIE Standard Illuminant D65, with a correlated color temperature of 6,500 K and a color-rendering index of 92. The visual angle subtended by the stimulus was about 2.5°. The colors were exposed using a solenoid-driven single-bladed shutter, and a computer was interfaced with the shutter mechanism to record response times.

Response times were recorded and used to distinguish "landmark" and "other basic" colors as was done in B&O. The same definition of *focal* colors was used: foci, which are the "best" examples of individual colour categories, were defined as those samples, *named with consensus*, that exhibited the shortest response time within their categories. Centroids, which represent the central location of colors within the categories, were determined according to a predefined metric.

Rationale

In this section, we provide the rationale for the experimental setting referring to the most relevant state-of-the-art contributions to the field.

There are some differences in the proposed set-up with respect to the B&O. First, the monolexemic color names were constrained to the 11 basic color categories identified by Berlin and Kay. The proposed work takes Berlin and Kay's findings as the starting point and rests on the assumption that at least 11 basic color categories do exist in languages at the seventh evolutionary stage. As mentioned above, in our view, while there is an evidence that focal colors tend to cluster in specific regions of the color space, as stated by the universalist hypothesis, the between-cultural variability can be accounted for by allowing a suitable level of fuzziness in the definition of the category boundaries, as predicted by the relativistic hypothesis. In this way, the 11 basic names hypothesis can be conciliated with the inter-subject and inter-cultural variability which shape our experiments.

In this work, we are not addressing the inter-cultural issue in general. We are considering the Italian and English languages taking the last one as the reference. We are aware that the comparison of our results with B&O and S&W would also require performing the experiments with native English speakers. This would allow checking the cross-cultural coherence assumption, namely if cross-linguistic differences do have any impact on color categorization. However, we assume here that between-cultural differences are comparable to within group variations (Kuehni, 2005; Webster et al., 2002).

The second difference between the two set-ups is in the observation conditions. In our experiment, colors are observed binocularly on a calibrated CRT monitor with a mid-gray background in a dark room. The question that we address is to what extent the rendering method impacts on the categorization process. This allows characterizing color naming within the state-of-the-art experimental equipments besides making our results suitable for further use in real applications (i.e. for designing color-naming models). On top of this, rendering the colors on a calibrated CRT

monitor in a dark room allows a strict control of the experimental conditions, improving the repeatability of the experiment and the reproducibility and accuracy of the data. A first work employing a CRT monitor was that of Smallman and Boynton (Smallman and Boynton 1990) who presented a series of experiments which sought to determine the extent to which optimal examples of basic colors segregate in an information display.

The third difference between the B&O experiment and ours is in the definition of focal colors. In Boynton, Schafer, and Neun (1964) and Sturges and Whitfield (1997), the focal colors were selected only for basic categories. In the original experiment, response times were recorded and used to classify the color names as basic or non-basic and focal colors were defined as the basic colors named *with consensus* featuring the shortest response time in the respective category. This leads to two main observations. First, this rests on the assumption that response times are robust indicators of saliency in this task. While B&O and S&W are in favor of this thesis, other researchers raised some concerns. In Alvarado and Jameson (2005), it is observed that fast response times tend to accompany higher confidence and accuracy; however, longer response times are ambiguous with respect to accuracy and confidence.

Very recently, Kuehni (2005) investigated the within-subject variability of foci based on the world color survey (WCS) data. Focal Hue (FU) range variability for Hering primaries (red, green, blue, yellow) was compared to Unique Hue (UH) range variability for such colors. The author concluded that there is a strong evidence that, with few exceptions, the focal color ranges (FCR) line up quite closely with the UH ranges. According to the author, this would support the hypothesis of a neurophysiological basis for the perceptual mechanisms guiding foci localization, that would dominate over cultural influences. These findings are in agreement with those of Webster et al. (2002). In their experiment, they addressed the question whether focal colors can differ systematically across different groups of individuals living in potentially different color environments. Consistent with many previous reports, the range of unique hues within individual groups was large, suggesting that the hue loci are very malleable. These results may reflect an influence of individual, environmental or cultural differences in focal color choices which, however, would represent a second order effect with respect to the prevailing mechanisms making between- and within-group variability very similar.

Our experiment on focal colors was inspired by these findings. Focal colors were selected by the subjects from among a set of stimuli as the *best representatives* of the respective category. In the identification of a unique (global) focal color per category, within- and between-group variability were considered as physiologic. Within-group variability was accounted for by summing the scores

given by the different subjects to the chosen focal color and taking the one with the highest score as the winner, as detailed in [Methods](#).

Finally, the choice of the focal colors was constrained to the consensus colors. The justification for this is twofold. First, both B&O and S&W identified foci as those samples, *named with consensus*, that exhibited the shortest response time within their categories. This implies the implicit constraint to the consensus set. Were a color not belonging to this sample set be named with the shortest response time, it would not have been considered eligible as a focal. It is reasonable to expect that, since the subjects were not informed, making it a prior instead of a posterior constraint would have left the resulting set of foci unchanged. Accordingly, we decided to constrain the set to consensus colors a priori. In consequence, we consider our results comparable with B&O's despite the choice of using constrained naming. However, since the task for B&O's subjects was unconstrained monolexic color naming, a straightforward comparison is only possible for subject CXO, who decided to constrain the name set to the 11 basic colors.

In our experiment, we used a different color reference set: the Extended OSA (E-OSA) color system (Menegaz et al., 2007; Menegaz, Le Troter, Boi, & Sequira, 2008). As mentioned above, other differences with respect to the B&O experiment are the rendering method, the use of unconstrained naming and the way the focal colors are extracted from the set of *consensus* colors. For the sake of comparison with B&O and S&W, results are presented using the same format and terminology as much as possible and taking the B&O data as reference. With respect to S&W, a larger discrepancy is expected because of the used of a different color model (Munsell).

Methods

Two experiments were performed, hereafter Exp1 and Exp2. The first one is the color-naming experiment, while the second one targets the identification of the focal colors. With respect to the identification of focal colors, our paradigm is different from that of B&O, whose definition of focal colors was based on the response time, as illustrated in the previous section.

The color samples were obtained by a topology-preserving resampling of the OSA color order system (Menegaz et al., 2008), leading to the E-OSA bases. The main features of the resulting color system are summarized in the next section.

Extended OSA-UCS basis

In OSA-UCS (Wyszecki & Stiles, 2000) color samples are arranged in a regular rhombohedral lattice in which each

color is surrounded by 12 neighboring colors, all *perceptually* equidistant from the considered one with good approximation. The OSA-UCS system has the unique advantage of equal supra-threshold perceptual spacing among color samples. However, the fact that the volume of the color space corresponding to the OSA samples fails to extend to highly saturated regions is a major limitation. Figure 1 illustrates the positions of the OSA samples in the CIELAB and sRGB spaces, respectively. The Extended-OSA (E-OSA) color basis was obtained by resampling the OSA space from the volume enclosing the original samples. The sampling was such that the properties of the coordinates of the specimens were preserved: L takes only integer values and $\{j, g\}$ are either both even or both odd. In this way, the topology of the OSA-UCS space was preserved. However, this does not guarantee that the new samples feature the same perceptual property of the original ones. This issue is left for future investigation. In this way, the total number of samples was increased to 1,104. The E-OSA samples falling out of the sRGB cube were excluded since they would correspond to colors that cannot be rendered on a monitor. Figure 2a and b illustrates the positions of the E-OSA samples in CIELAB and sRGB spaces, respectively. The comparison with Fig. 1 reveals an extended gamut with respect to OSA which covers the sRGB cube almost entirely. The regions that still remain partly uncovered are those close to the edges of the sRGB cube, especially those corresponding to high luminance values (i.e. the highly saturated yellows). The E-OSA basis is not exactly a superset of the B&O set, as 20 of the 424 samples are missing. This is due to the fact that the sRGB coordinates of such colors would fall outside the sRGB cube and thus were excluded.

Experiment 1: color naming

A psychophysical experiment was conducted for the categorization of the E-OSA color samples. The CRS ViSaGe¹ board for high quality stimuli presentation was used (Menegaz et al., 2008). Colors were rendered on a calibrated Toshiba Diamond Pro 2070 monitor in a completely dark room. Each color sample was shown in a square window of size 2×2 cm² subtending a visual angle of 2° in a mid-luminance gray background. Viewing distance was set to 57 cm. Samples were presented one at a time in random order. The order was different for each block of trials and within trials for the same subject. The task consisted in assigning each color sample to one of the 11 categories by selecting one among eleven keys mapped to the numerical keys on the keyboard. No time limit was

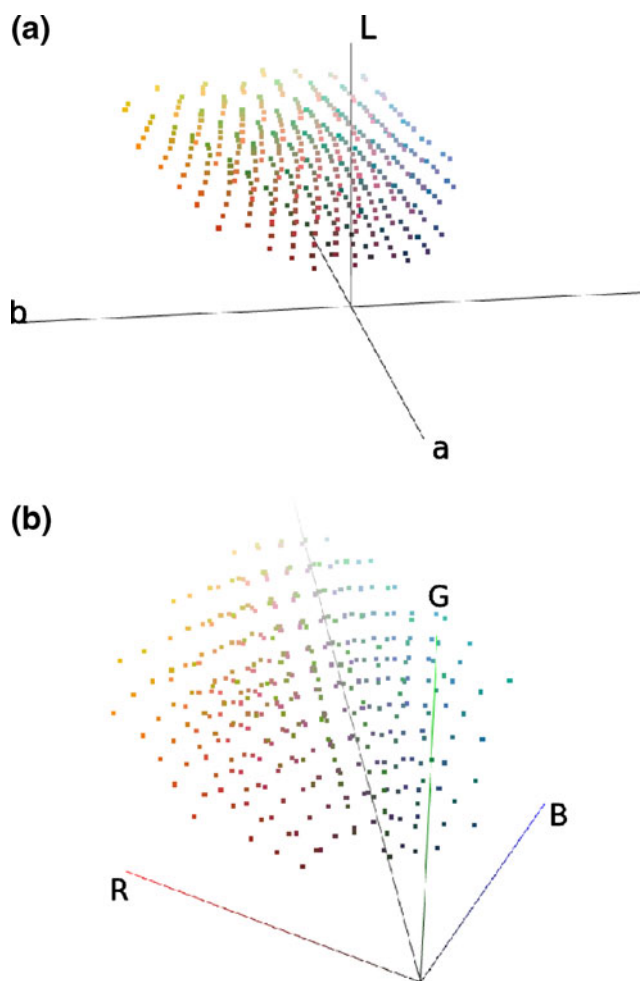


Fig. 1 The 424 OSA-UCS samples represented in **a** CIELAB and **b** sRGB spaces

set. In order to avoid after-images a mid gray luminance image (the same used as background) was interleaved to the stimuli for a duration of 500 ms. Subjects were asked to provide an answer within the shortest delay without sacrificing accuracy.

Each trial was split in two blocks consisting of 501 and 513 color samples, respectively, in order to avoid the decay of attention due to fatigue. Each block of trials required twenty minutes on average.

Seven Italian native speakers aged between 25 and 40 participated in the experiment. Out of these, 5 performed it twice.

All subjects had normal or corrected-to-normal vision. Subjects were screened for color vision using the Ishihara test.

Experiment 2: focal colors

As discussed in the Rationale, the focal colors were constrained to the set of consensus colors. The test colors were rendered as in Exp1 and observed under the same

¹ <http://www.crs ltd.com/>

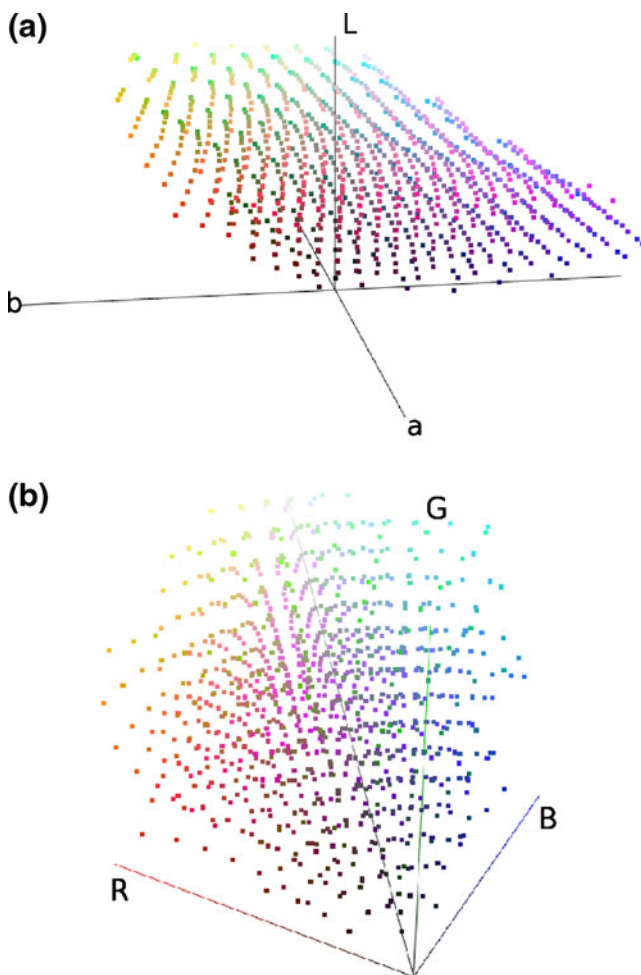


Fig. 2 The E-OSA samples represented in **a** CIELAB and **b** sRGB spaces

viewing conditions to guarantee experimental uniformity. Stimuli were displayed one at a time on the same monitor as in Exp1 and each stimulus were subtending a visual angle of 2 degrees. No limit to the response time was given. For each stimulus, the subjects were asked to score in a number between zero and 100 the perceived closeness of the sample to the category prototype, where a score of 100 corresponded to perceptual correspondence. The current class was indicated in characters on top of the screen. The values were input by the keyboard. Response times were recorded. In order to limit the fluctuations in response time due to the gaze direction, a fixation point was set at the center of the screen. A mid-gray uniform image was displayed between successive stimuli presentations to avoid after images for a duration of 500 ms. Seven Italian native speakers participated in the experiment. Of these, six also participated in Exp1. A total of 357 colors were used, distributed as follows: 13 red, 80 green, 91 blue, 4 yellow, 19 orange, 60 purple, 80 pink and 10 brown. No consensus colors were identified in the achromatic color categories.

The individual scores were added for each test color and the focal colors were defined as those having the highest score in each category.

Results

Experiment 1

Consensus, consistency and centroids were determined and compared to those obtained by B&O and S&W. The following parameters were considered: consistency rate; position of consensus colors; positions of centroids. All these comparisons were performed both in the OSA and in the CIELAB color systems, featuring perceptual uniformity. Furthermore, the distance among the focal colors and the centroids within and between color categories were analyzed. The following color systems were considered: B&O, E-OSA, the set of 404 colors that represents the overlap between E-OSA and B&O (hereafter called Reduced(R)-OSA), and S&W.

As mentioned above, we regard an agreement of color naming by a single subject for two presentations as defining consistency. Consensus is reached when all subjects name a color sample consistently using the same basic color term.

The centroids was determined by averaging the $\{L, j, g\}$ values for all samples assigned to a particular name, weighted according to whether the name was used once or twice. Calculations were performed for the sake of comparison with B&O.

Consistency

Figure 3 shows the rate of consistently named colors for the four bases. Since the number of subjects is different in our and B&O experiments, the comparison can only be made on the consistency rates.

Data show that the green category gathers the largest consistency for B&O, R-OSA and S&W reaching the 30, 25 and 22%, respectively. The consistency rate is nearly equal in the three samples B&O, R-OSA and S&W for the blue category (17.98, 17.38, and 17.40%). Blue is the one that gathers the largest consistency for E-OSA (23.27%) while green reaches 17.58%. Except for the achromatic colors, red has small consistency rate in all samples (3.89% for B&O, 3.32% for R-OSA, 4.42% for E-OSA and 6.11% for S&W). Furthermore, yellow and brown for E-OSA and orange for S&W are below the consistency rate of red. A large variation in the consistency rate can be observed for pink (9–22%) and purple (5–16%) categories. A similar behavior can be found in the classes of yellow and orange, whose consistency rates vary between 3 and 9.5% and between 5 and 12.5%, respectively. In the B&O sample,

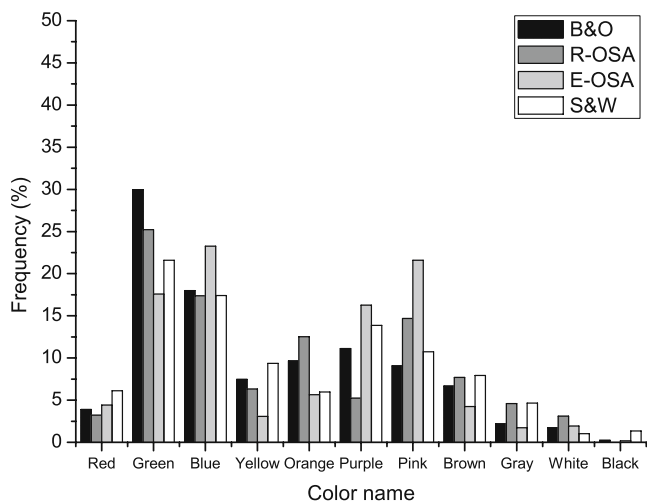


Fig. 3 Consistency rate for the four reference bases: B&O, R-OSA, E-OSA, S&W. This was evaluated as the percentage of consistently named samples per category with respect to the total number of colors named consistently, in each basis. B&O *black bars*; R-OSA *gray bars*; E-OSA *light gray bars*; S&W *white bars*

purple reaches a larger consistency than pink and orange, while the trend is opposite for R-OSA, for which pink is followed by orange and purple, in the order. E-OSA and S&W show similar behavior in this respect, having orange with smaller consistency than pink. For E-OSA, the largest consistency rate is reached by blue followed by pink, green and purple, in the order.

Consensus

The total number of consensus colors was 357, 112, 128 and 102 for the E-OSA, R-OSA, B&O and S&W samples, respectively.

Figure 4 presents the results for each of the considered basis. In particular, Fig. 4a shows the rate between the total number of consensus colors per class and the total number of consensus colors; Fig. 4b gives the rate between the total number of consensus colors per class and the total number of samples in the basis. Figure 5 provides the percentage of consensus colors in each basis.

For B&O and R-OSA the category featuring the largest consensus is green, followed by blue. For E-OSA, this trend is inverted as blue is followed by green and pink, with same rate between the total number of consensus colors per class and the total number of consensus colors. The same trend for green and blue can be observed in S&W, with the difference that pink has a lower rate than green.

Consensus on achromatic colors was not reached in E-OSA, thus in R-OSA, and for B&O, no consensus was reported for black. A large rate of variability was observed for yellow, orange, pink and purple over the different bases, as was the case for consistency. The consensus rate with

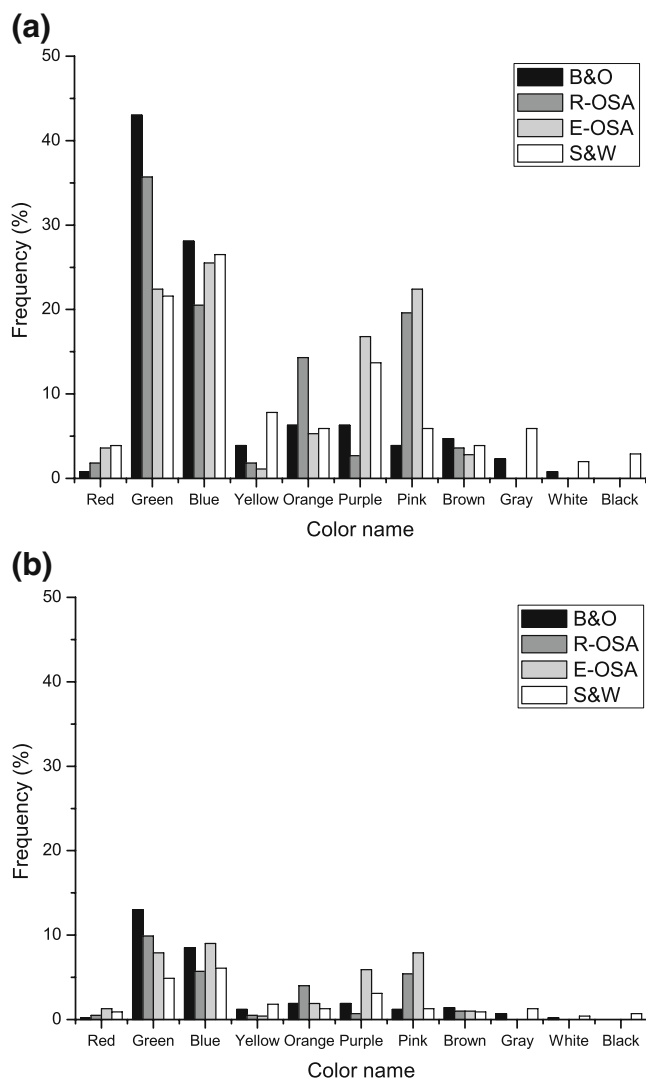


Fig. 4 For each basis: **a** rate between the total number of consensus colors per class and the total number of consensus colors; **b** rate between the total number of consensus colors per class and total number of samples in the basis. B&O *black bars*; R-OSA *gray bars*; E-OSA *light gray bars*; S&W *white bars*

respect to the numerosness of the sample was the highest for E-OSA (35%), followed by B&O (30%), R-OSA (28%) and S&W (23%).

Agreement on consensus colors was reached on 64 color samples. Tables 1 and 2 show the distribution of consensus samples per class and per lightness level for the R-OSA and B&O bases, respectively. While green and blue can be found in the quasi-totality of the plans, as it is reasonable to expect, it is quite surprising that pink spans over $L = -2, \dots, 4$ in the R-OSA bases. No consensus was reached in R-OSA for achromatic colors, which is also the case for black in B&O bases. At very high lightness levels, no consensus was reported for R-OSA except for one sample in the pink class. Conversely, samples for the classes of yellow, blue and white reached consensus at high lightness levels for B&O. A

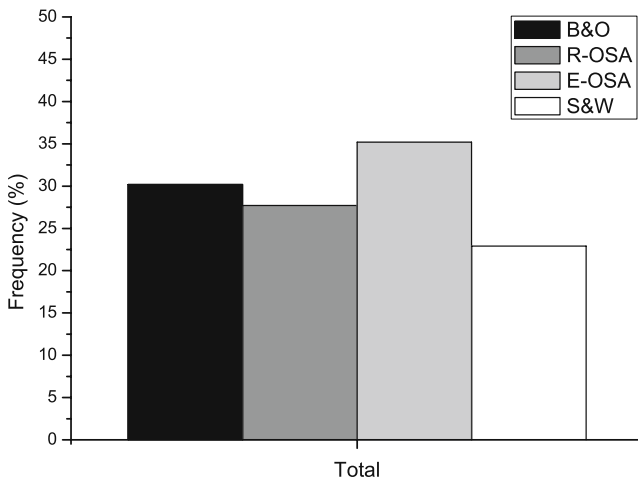


Fig. 5 For each basis: rate of consensus samples per basis. B&O black bars; R-OSA gray bars; E-OSA light gray bars; S&W white bars

similar trend was detected for very low lightness levels. However, the very small number of consensus colors for both bases only supports qualitative considerations. A particular case is that of $L = -4$. In this plan, B&O found many green, blue and purple consensus samples, one red, two brown and one gray, while only two consensus colors for green and one for red were found in R-OSA. The last row shows the percentage of consensus colors per lightness value with respect to the total number of samples having the same lightness. The last column provides the percentage of consensus colors per color class. Interestingly, the pink category gathers large consensus in the R-OSA basis, compared to that of blue, while achromatic colors are very poorly represented.

Centroids

Centroids were found by taking the weighted average of the coordinates of all colors in the bases, the weights w_j^i being the number of times the color was assigned to a given class by all subjects

$$w_j^i = \frac{1}{N_{trials}} \sum_{k=1}^{N_{trials}} v_k(j, i) \tag{1}$$

where w_j^i is the weight given to color j ($j = 1, \dots, 1014$) when evaluating the centroid of class i ($i = 1 \dots 11$), N_{trials} is the number of trials ($N_{trials} = 12$) and $v_k(j, i)$ is equal to either one or zero depending on the fact that color j was assigned to class i in trial k or not. The centroid coordinates for the i^{th} category become

$$L^i = \sum_{j=1}^{1014} w_j^i L_j$$

$$j^i = \sum_{j=1}^{1014} w_j^i j_j$$

$$g^i = \sum_{j=1}^{1014} w_j^i g_j \tag{2}$$

Table 3 shows the $\{L, j, g\}$ coordinates of the centroids for the B&O, R-OSA, E-OSA and S&W sets. For the last one, the Munsell coordinates of the centroids were converted to $\{L, j, g\}$ by taking the $\{L, j, g\}$ coordinates of the samples that is the closest to the centroid in the Munsell space.

Table 1 Distribution of R-OSA consensus samples as a function of category and lightness

	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	Tot	% Tot
Red	0	0	0	1	1	0	0	0	0	0	0	0	0	2	1.7
Green	1	2	4	2	6	5	7	4	5	4	0	0	0	40	35.7
Blue	1	2	1	0	4	3	3	2	3	3	1	0	0	23	20.5
Yellow	0	0	0	0	0	0	0	0	0	0	2	0	0	2	1.7
Orange	0	0	0	0	0	1	4	3	6	2	0	0	0	16	14.2
Purple	0	1	1	0	1	0	0	0	0	0	0	0	0	3	2.6
Pink	0	0	0	0	0	4	3	3	5	2	4	1	0	22	19.6
Brown	1	1	2	0	0	0	0	0	0	0	0	0	0	4	3.5
Gray	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
White	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Black	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N_{cons}	3	6	8	3	12	13	17	12	19	11	7	1	0	112	100
N	6	17	22	27	36	43	46	49	48	45	34	23	8	404	
%	50	35.2	36.3	11.1	33.3	30.2	36.9	24.4	39.5	24.4	20.5	4.3	0	27.7	

N_{cons} and N are the total number of consensus colors and total number of color samples, respectively, per each value of L . Last row represents the rate N_{cons}/N as a percentage

Table 2 Distribution of B&O consensus samples as a function of category and lightness

Lightness Level															
	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	Tot	% Tot
Red	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0.7
Green	1	2	3	4	6	9	7	9	8	4	1	1	0	55	42.9
Blue	0	2	3	2	4	5	4	5	5	3	2	1	0	36	28.1
Yellow	0	0	0	0	0	0	0	0	0	0	2	3	0	5	3.9
Orange	0	0	0	0	0	0	2	4	2	0	0	0	0	8	6.2
Purple	1	1	2	3	1	0	0	0	0	0	0	0	0	8	6.2
Pink	0	0	0	0	0	0	0	2	1	1	1	0	0	5	3.9
Brown	0	1	3	2	0	0	0	0	0	0	0	0	0	6	4.6
Gray	0	0	0	1	0	1	0	1	0	0	0	0	0	3	2.3
White	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0.7
Black	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N_{cons}	2	6	11	13	11	15	13	21	16	8	6	6	0	128	100
N	6	17	24	29	39	47	47	49	48	45	37	28	8	424	
%	33.3	35.2	45.8	44.8	28.2	31.9	27.6	42.8	33.3	17.7	16.2	21.4	0	30.1	

N_{cons} and N are the total number of consensus colors and total number of color samples, respectively, per each value of L . Last row represents the rate N_{cons}/N as a percentage

It can be observed that the centroid for blue for E-OSA has a lower lightness than in B&O while the opposite happens for green. The lightness values of the centroids of all classes for S&W are quite different from those of the other sample sets. This is most probably due to the different color model taken as reference. Results show a good overlap among the centroids for all the categories but the yellow, for which B&O reports a higher lightness and green-blue component. Figure 6 gives the distances between the centroids of the color categories of the B&O

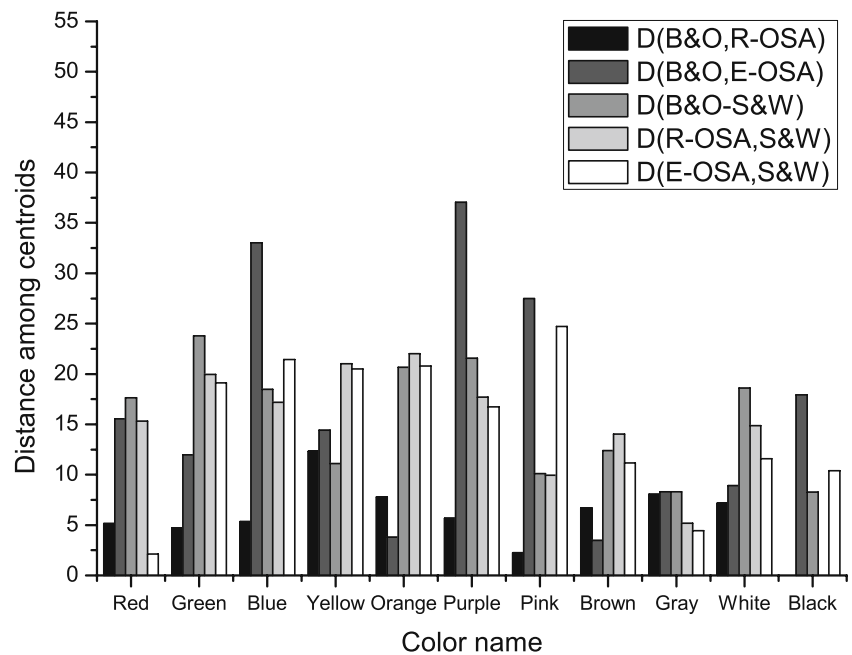
basis with respect to the R-OSA, E-OSA and S&W basis, respectively, as well as the distances between the centroids of E-OSA and R-OSA to S&W, respectively. The Euclidean metric in the CIELAB color model was used.

The distances between centroids of B&O and R-OSA are systematically smaller than those between B&O and E-OSA with the exception of brown and orange. The largest distance $D(B&O,R-OSA)$ corresponds to yellow (12.32) and the largest distance $D(B&O,E-OSA)$ corresponds to purple (37.05). The second largest $D(B&O,E-OSA)$ corresponds to

Table 3 $\{L, j, g\}$ coordinates of the centroids of each class for the four reference sample sets. The B&O formula (2) was applied

Centroids												
color name	B&O			R-OSA			E-OSA			S&W		
	L	j	g	L	j	g	L	j	g	L	j	g
Red	-3.6	1.5	-6.9	-3.5	2.2	-6.9	-3.1	2.4	-9.6	-3.05	2.52	-10
Green	-0.4	3.9	2.5	-0.8	3.4	2.8	0.4	4.6	4.2	-2.21	2.77	5.06
Blue	-0.8	-2.5	2.8	-0.9	-2.8	2.1	-1.3	-6.7	1.2	-1.52	-5.40	3.92
Yellow	2.6	8.2	-1.7	2.0	7.2	-0.2	2.6	7.6	0.4	3.57	9.50	-1.58
Orange	-0.2	5.7	-6.3	-0.1	5.7	-4.9	-0.1	5.7	-5.7	0.52	8.45	-5.68
Purple	-2.5	-2.5	-1.8	-3.1	-2.6	-2.5	-2.5	-6.9	-5.4	-3.34	-4.97	-3.60
Pink	0.5	0.3	-4.9	0.1	0.6	-4.8	0.1	-3.0	-8.6	0.51	0.84	-6.71
Brown	-3.2	2.8	-2.9	-2.3	3.1	-2.4	-3.6	2.9	-2.5	-3.85	4.29	-2.71
Gray	-1.3	0.3	0.1	-1.1	-0.8	-0.4	-1.2	-0.9	-0.4	-2.00	-0.62	-0.41
White	4.0	1.7	-0.3	3.7	0.6	-0.1	4.3	0.3	0.1	5.65	-1.07	-0.71
Black	-6.2	-0.5	0.5	-	-	-	-9.1	-0.3	-0.2	-7.35	-0.29	-0.18

Fig. 6 Euclidean distance between the category centroids of: B&O and R-OSA, E-OSA and S&W; the R-OSA and S&W; the E-OSA and S&W, respectively. The distances were evaluated in the CIELAB color model. $D(B\&O,R-OSA)$ black bars; $D(B\&O,E-OSA)$ dark gray bars; $D(B\&O,S\&W)$ gray bars; $D(R-OSA,S\&W)$ light gray bars; $D(E-OSA,S\&W)$ white bars



blue (33.03) followed by pink (27.50). The centroid distance between B&O and S&W is maximised for green, followed by orange and purple. $D(E-OSA,S\&W)$ is quite large for all classes except red (2.14).

Experiment 2 (focal colors)

Focal colors were identified via Exp2. This was performed based on the consensus colors resulting from Exp1. Even

though, for the sake of comparison with B&O, R-OSA basis would have been the most natural choice, we decided to use the wider sample E-OSA to keep more in line with our concept of focal colors being prototypical for the corresponding class. Figure 7 shows the positions of the centroids (empty symbols) and foci (filled symbols) for all the categories in the CIELAB space. B&O data are represented as squares, E-OSA as circles and S&W as diamonds, respectively.

Fig. 7 Focal colors and centroids for E-OSA (circles), B&O (squares) and S&W (diamonds) after projection to the (a,b) plan of the CIELAB space. Filled symbols correspond to focal colors

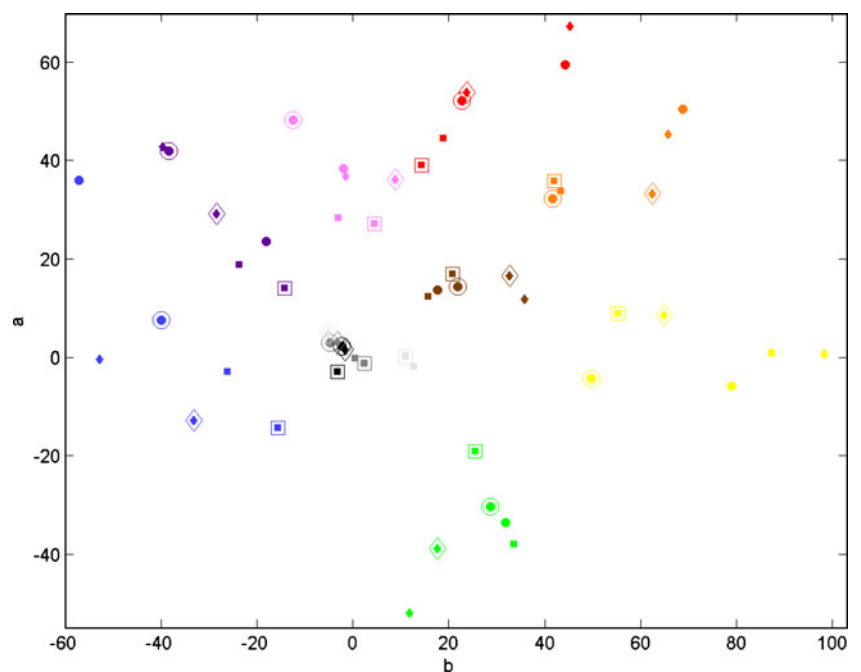


Table 4 $\{L, j, g\}$ coordinates of focal colors for B&O, E-OSA and S&W basis

Focal colors									
Name	E-OSA			B&O			S&W		
	L	j	g	L	j	g	L	j	g
Red	-3	5	-11	-4	2	-8	-2.49	5.01	-12.91
Green	-6	4	4	-1	5	5	-3.77	1.90	6.42
Blue	-7	-9	1	-6	-4	2	-1.18	-8.87	4.25
Yellow	3	11	1	4	12	0	3.52	12.73	0.2
Orange	1	9	-9	0	6	-6	0.88	8.74	-7.99
Purple	-7	-3	-3	-4	-4	-2	-3.14	-7.07	-5.35
Pink	3	-1	-7	3	-1	-5	2.42	-0.87	-6.68
Brown	-8	2	-2	-6	2	-2	-5.65	4.38	-1.73
Gray	-	-	-	-2	0	0	-2.00	-0.62	-0.41
White	-	-	-	4	4	0	5.65	-1.07	-0.71

Table 4 provides the coordinates of the foci as determined by Exp2, as well as those of the focal colors identified by B&O and S&W, respectively. All the B&O focals have higher lightness than E-OSA ones except for red and orange. S&W's focals have lightness levels above B&O ones except for pink and green.

Centroids were determined using formula (2) in $\{L^i, j^i, g^i\}$ and the resulting coordinates were transformed to CIELAB following the method proposed in Cao, Pokorny, & Smith, (2005).

Figure 8 gives the distances among the focal colors and the centroids for each class for B&O, E-OSA and S&W datasets, while Fig. 9 illustrates the distances

among the focal colors of each class between (1) B&O and E-OSA that is $D(B\&O, E-OSA)$; (2) E-OSA and S&W that is $D(E-OSA, S\&W)$, and (3) B&O and S&W that is $D(B\&O, S\&W)$, respectively.

It is interesting to note that the distance is quite large for the blue, green and purple categories for the E-OSA sample set. This could be an indication of the fact that those classes are too wide to be associated to a single name in the Italian language. The distance is the largest for the class of blue, where it reaches 47.28 units. The focal for blue corresponds to a lightness value much smaller than that of the centroid, which could be an indication of the fact that an additional term in the medium-to-high lightness region is needed for this category.

Fig. 8 Distance among the focal colors and the centroids per class for the B&O, E-OSA and S&W datasets. In this case, the centroids were evaluated as in formula (2) in $\{L, j, g\}$ and then moved to CIELAB. E-OSA black bars; B&O gray bars; S&W white bars

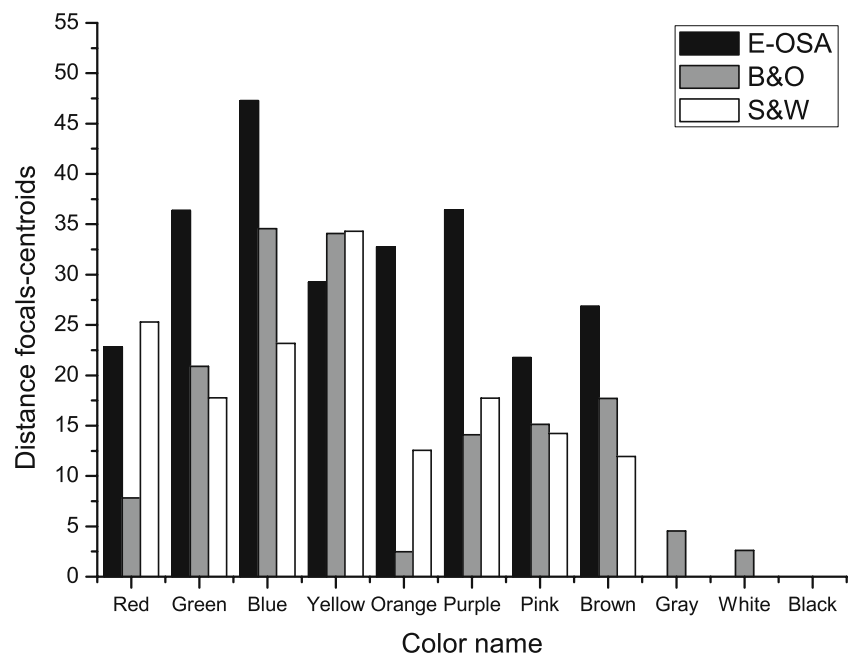
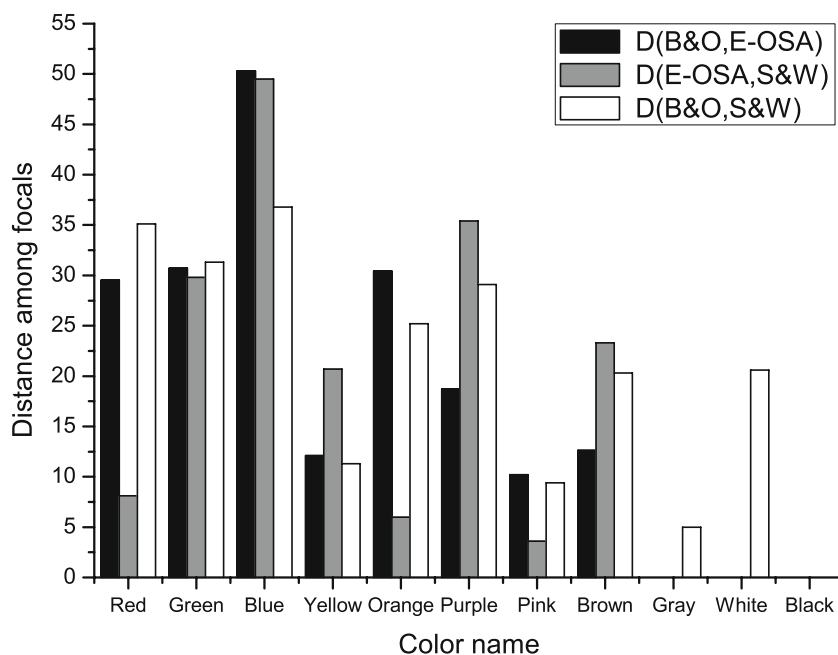


Fig. 9 Distance among the focal colors per class for the three datasets E-OSA, B&O and S&W. $D(B\&O, E-OSA)$ black bars; $D(E-OSA, S\&W)$ gray bars; $D(B\&O, S\&W)$ white bars



The distance between foci and centroids is interesting for investigating the central tendency of the color categories. For E-OSA, a very large distance was observed for blue (47.28). The fact that this is mostly due to a difference in lightness ($\{L, j, g\}_{centroid} = \{-1.3, -6.7, 1.2\}$, $\{L, j, g\}_{focal} = \{-7, -9, 1\}$) could suggest that colors in the blue category tend to cluster in a region that is not representative for the prototypical example of the category. This could be an indication that another class is needed in the light blue region, as it is the case for Greek, Russian and Turkish (Androulaki et al., 2006; Özgen & Davies, 1998; Paramei, 2005; Winawer et al., 2007). This trend was also observed for B&O's blue ($\{L, j, g\}_{centroid} = \{-0.8, -2.5, 2.8\}$, $\{L, j, g\}_{focal} = \{-6, -4, 2\}$). However, in this case, the distance is smaller (34.55) and comparable to that observed for green (36.38), purple (36.43) and orange (32.76) for E-OSA, and yellow for B&O (34.08) and for S&W (34.30). It is our opinion that the case of blue is peculiar, and we believe that further investigation is needed in this respect for the Italian language.

In E-OSA, the next largest distances among foci and centroids are observed for purple and green followed by orange. In B&O, the order is blue, yellow and green, while it is yellow followed by red, green and purple for S&W.

Conclusions

Two experiments were conducted. The first experiment (Exp1) was a constrained color-naming experiment. Both the E-OSA and the reduced R-OSA samples were considered. The test was performed by Italian native

speakers and colors were observed binocularly on a calibrated CRT monitor under controlled conditions. The second experiment (Exp2) was conducted for identifying focal colors. A different paradigm with respect to B&O was followed even though response times were recorded. The same formula as B&O was used to determine the centroid coordinates in $\{L, j, g\}$. Comparison were also performed with S&W data. Results show some similarity among the data obtained by B&O with R-OSA while a less pronounced agreement is observed, in general, with the E-OSA sample. The comparison with S&W implies an additional difference, as in that case the Munsell sample was used.

Overall, results show that green and blue span over a very large set of lightness levels while yellow, purple and orange are confined to smaller portions of the color space. The largest distance among foci and centroids corresponds to the blue category for both E-OSA and B&O. However, in the case of E-OSA, such a distance reaches a value that is 10 units above the maximum distance recorder for all the other cases. This could be an indication of the fact that an additional class corresponding to light-blue could be needed in the Italian language, as it is the case for Greek, Russian and Turkish. This is currently under investigation.

In summary, results show a difference in the partitioning of the colorspace. Furthermore, it was observed that achromatic colors are still missing in E-OSA and would need to be included for future studies. Other experiments using a suitable reference basis are currently being performed, including the use of the Munsell basis.

References

- Alvarado, N., & Jameson, K. (2005). Confidence judgments on color category best exemplars. *Cross-Cultural Research*, 39(2), 134–158.
- Androulaki, A., Gómez-Pestaña, N., Mitsakis, C., Jover, J., Coventry, K., & Davies, I. (2006). Basic colour terms in modern Greek: Twelve terms including two blues. *Journal of Greek Linguistics*, 7(1), 3–47.
- Benavente, R., Vanrell, M., & Baldrich, R. (2008). Parametric fuzzy sets for automatic color naming. *Journal of the Optical Society of America A*, 25(10), 2582–2593.
- Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley: University of California.
- Bleys, J. (2004). *The cultural propagation of color categories: Insights from computational modelling*, (PhD thesis). Vrije University Brussel: Brussels.
- Boynton, R., & Olson, C. (1987). Locating basic colors in the OSA space. *Color Research & Application*, 12(2), 94–105.
- Boynton, R., Schafer, W., & Neun, M. (1964). Hue-wavelength relation measured by color-naming method for three retinal locations. *Science*, 146(3644), 666.
- Cao, D., Pokorny, J., & Smith, V. (2005). Associating color appearance with the cone chromaticity space. *Vision Research*, 45(15), 1929–1934.
- Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a stone-age tribe. *Nature*, 398(6724), 203–204.
- Dedrick, D. (1997). Colour categorization and the space between perception and language. *Behavioral and Brain Sciences*, 20(02), 187–188.
- Hardin, C. (2005). Explaining basic color categories. *Cross-Cultural Research*, 39(1), 72–87.
- Jameson, K. (2005). Why GRUE? An interpoint–distance model analysis of composite color categories. *Cross-Cultural Research*, 39(2), 159–194.
- Jameson, K., & D’Andrade, R. (1997). Color categories in thought and language. In C. L. Hardin (Ed.), (pp. 295–319). Cambridge: Cambridge University Press.
- Jameson, K., & Alvarado, N. (2003). The relational correspondence between category exemplars and names. *Philosophical Psychology*, 16(1), 25–49.
- Kay, P., & Kempton, W. (1984). What Is the Sapir-Whorf Hypothesis? *American Anthropologist*, 86(1), 65–79.
- Kay, P., & Regier, T. (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences of the United States of America*, 100(15), 9085.
- Kay, P., Berlin, B., Maffi, L., & Merrifield, W. (1997). Color naming across languages. *Color categories in thought and language*, 21.
- Kuehni, R. (2005). Focal color variability and unique hue stimulus variability. *Journal of Cognition and Culture*, 3(4), 409–426.
- Lammens, J. ele. (1994). *A computational model of color perception and color naming* (Unpublished doctoral dissertation). Citeseer.
- Lin, H., Luo, M., MacDonald, L., & Tarrant, A. (2001c). A cross-cultural colour-naming study. part i: Using an unconstrained method. *Color Research and Applications*, 26, 40–60.
- Lin, H., Luo, M., MacDonald, L., & Tarrant, A. (2001b). A cross-cultural colour-naming study. part ii: Using a constrained method. *Color Research and Applications*, 26, 193–208.
- Lin, H., Luo, M., MacDonald, L., & Tarrant, A. (2001a). A cross-cultural colour-naming study. Part III—A colour-naming model. *Color Research & Application*, 26(4), 270–277.
- Man, T., & MacAdam, D. (1989). Three-dimensional scaling of the uniform color scales of the Optical Society of America. *Journal of the Optical Society of America A*, 6(1), 128–138.
- Menegaz, G., Le Troter, A., Sequeira, J., & Boi, J. (2007). A discrete model for color naming. *EURASIP Journal on Applied Signal Processing*, 2007(1), 113.
- Menegaz, G., Le Troter, A., Boi, J., & Sequeira, J. (2008). Semantics driven resampling of the osa-ucs. In *Image analysis and processing workshops, 2007. iciapw 2007. 14th international conference on* (pp. 216–220).
- Özgen, E., & Davies, I. (1998). Turkish color terms: Tests of Berlin and Kay’s theory of color universals and linguistic relativity. *Linguistics*, 36(5), 919–956.
- Paramei, G. (2005). Singing the Russian blues: An argument for culturally basic color terms. *Cross-Cultural Research*, 39(1), 10–34.
- Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences*, 104(4), 1436.
- Roberson, D. (2005). Color categories are culturally diverse in cognition as well as in language. *Cross-Cultural Research*, 39(1), 56–71.
- Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology General*, 129(3), 369–398.
- Saunders, B., & Van Brakel, J. (1997). Are there nontrivial constraints on colour categorization? *Behavioral and Brain Sciences*, 20(02), 167–179.
- Smallman, H., & Boynton, R. (1990). Segregation of basic colors in an information display. *Journal of the Optical Society of America A*, 7(10), 1985–1994.
- Sturges, J., & Whitfield, T. (1997). Salient features of Munsell colour space as a function of monolexic naming and response latencies. *Vision Research*, 37(3), 307–313.
- Webster, M., Webster, S., Bharadwaj, S., Verma, R., Jaikumar, J., Madan, G., et al. (2002). Variations in normal color vision. III. Unique hues in Indian and United States observers. *JOSA A*, 19(10), 1951–1962.
- Winawer, J., Witthoft, N., Frank, M., Wu, L., Wade, A., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, 104(19), 7780.
- Wyszecki, G., & Stiles, W. (2000). *Color science: Concepts and methods, quantitative data and formulae*, 2nd ed. Wiley-VCH