

Categorical Formation of Mandarin Color Terms at Different Luminance Levels

T. J. Hsieh,¹ I. P. Chen^{2*}

¹Department of Information Communication, Chinese Culture University, Taipei, Taiwan

²Institute of Applied Art, National Chiao Tung University, Hsin-Chu, Taiwan

Received 8 October 2009; revised 30 January 2010; accepted 5 May 2010

Abstract: This study presents the categorical formation of a set of Mandarin color terms on the International Commission on Illumination (CIE) 1931 chromaticity diagram across six luminance levels. This article conducted a study that employed 44 native Mandarin speakers to perform a force-choice sorting task. The Mandarin color terms for sorting were determined by a free-recall pretest and are consistent with basic color terms proposed by Berlin and Kay. The square-sampled stimuli were generated by evenly sweeping the x - y diagram of 5, 10, 25, 50, 100, and 170 cd/m^2 planes. The categorical sorting results and response time (RT) measurements suggest that: (1) the concepts of green, blue, purple, and gray stably exist at most luminance levels. The voting RT for the green, blue, and purple categories is particularly short. (2) Red, orange, yellow, and pink are highly luminance-dependent; these can be identified without difficulty only at some restricted luminance levels. (3) The chromaticity areas designated as orange, partial yellow, red, and pink are recognized as brown when the luminance level decreases. (4) Brown and gray serve as representations of two distinct tints in the low saturation condition. (5) The location of boundaries between blue and green are remarkably different than those in a similar study that employed Japanese speakers. © 2011 Wiley Periodicals, Inc. *Col Res Appl*, 36, 449–461, 2011; Published online 12 August 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/col.20638

Key words: color terms; categorical; Mandarin; CIE1931; luminance

INTRODUCTION

The mind is fundamentally able to process vast information categorically, and this many-to-one mechanism also functions in color perception. Various color experiences are induced by the combination of visible wavelengths with infinite possibilities, but the visual system treats this changing continuum discretely by specifically sorting color stimuli into various qualitative categories.¹ Between-category color discriminations are more accurate and efficient than equivalent within-category discriminations.^{2–4} Distinguishable color categories in the human perception system are limited in number; there are about 200 distinct steps,⁵ or 120–150 just noticeable differences¹ across the visible spectrum. This categorical effect is also apparent in color memorization.^{3,6–8}

The number of color categories at the linguistic level is even more restricted than at the visual discrimination level. Linguistic color categories can be designated by monolexemic terms like red, green, yellow, and blue.^{9,10} Color category terminology is an important issue in linguistic and cognitive science.^{11,12} Berlin and Kay established a pioneering theory of basic color terms by conducting an anthropological survey; they propose 11 common color terms that are widely used across cultures.¹³ These universal color terms in English are black, white, red, green, yellow, blue, orange, purple, pink, brown, and gray. This color sequence is based on the developmental order of the terms. This universalist view of the usage of color terms, involving a belief in profound common ground that connects human cultures and minds, has been observed in a variety of disciplinary studies, including cross-culture surveys,^{4,14–18} free color-naming tasks,^{9,10,17} developmental studies,^{19–22} and psychophysics and physiological experiments,^{9,23–30} despite the continued existence of opposing, relativist arguments.^{31–33} These two opposing stances in linguistic anthropology—universalist and relativist—engage in intense debate regarding the

*Correspondence to: I. P. Chen (e-mail: iping@faculty.nctu.edu.tw).

dominant hierarchy of language and thought.^{12,31,34} The universalist view holds that language is a limited semantic palette shaped and restricted by human cognition. The range of color categories are projected from the universal color foci and therefore located in similar positions in color space across world languages. In contrast, the relativist view denies the universal foci theory and argues that language shapes thought. According to this view, color categories are defined at their boundaries by local language conventions and may vary widely across cultures.³² Recently, a third theory has arisen, which claims that color naming reflects optimal or near-optimal divisions of the irregularly shaped perceptual color space.^{11,35} This hypothesis seems to be confirmed by tests of the hidden consensus of world color survey (WCS).³⁵

Color naming and categorization are topics of intense investigation in part because color is a salient visual feature in most human cultures. Additionally, the “thought” of color can be scientifically defined in the chromaticity space through standard measurement techniques. In other words, the appropriate experimental survey can convert the color semantics from the linguistic domain to the physical domain. To gauge the corresponding chromaticity range of the color term, two methods, free naming method (the unconstrained method³⁶) and sorting method (the constrained method³⁶), are frequently adopted. These two methods probe very different aspects in color naming and categorizing issue.

The free naming method can collect a large, diverse amount of color name data,^{9,10,17,18,37,38} whereas the sorting method specifically focuses on mapping the corresponding chromaticity range of the color terms in question.^{18,36,39} In a free naming task, participants are usually instructed to use freely monolexic color term^{9,37} to name the present color stimulus. The frequency counts of free naming results reveal the dominant color names used in a culture. For instance, Boynton and Olson used naming method to confirm the perceptual salience of the 11 basic color terms^{13,29} out of various color terms produced by their participants.

In a typical sorting task, also called the constrained method,³⁶ participants are given a set of color terms as options for classifying the presented color stimulus. Methodologically, the sorting task uses force-choice tasks and systematic stimuli sampling, which can efficiently bridge each color term and its corresponding area in the chromaticity coordinate. The color terms used would vary depending on the purpose of the study. Boynton *et al.* used four perceptually unique colors, red, green, yellow, and blue as dependent factors in accessing the perception of stimuli, and they determined the relation between the characteristics of color vision and these colors.^{25,28} The results indicated that the colors of the opponent color pair, red and green, and yellow and blue, were rarely identified together when describing the color stimulus. Besides the set of unique colors, the 11 basic color terms were widely adopted as sorting options in related studies due to their generality among human cultures.^{9,16,29}

Over decades of color-naming data collection, mapping basic color terms on the Munsell 330-colors palette across different languages has been executed and integrated. This comparative result is known as WCS.^{40,35} The remarkable cross-culture surveys provide empirical evidence for establishing psycho-anthropological theories regarding the relation between color perception and language development. However, the environmental lighting in the earlier studies was not finely controlled, and their stimuli were reflective materials, which increase the difficulty of generalizing the naming behavior in following color studies and applications. Although many recent studies specified stimuli on the standard color space, the detailed experimental settings are diverse and consequently the cross-comparison among their results seems to be inappropriate. For instance, Lin *et al.* adopted Inter Society Color Council- National Bureau of Standards (ICCS-NBS) samples on glossy papers as stimuli and used both free naming and sorting methods. Their color terms for sorting were adopted from a previous free naming task instead of known basic color terms.³⁶ Guest and Larr¹⁰ used cathode ray tube (CRT)-displayed stimuli and free naming method and plotted the naming responses on Lu'v' space. Their study found a significant pattern in the behaviors of color naming, which reveals the forming structure of color categories within individual. Shinoda *et al.* used CRT-displayed, regular-sampled stimuli and sorting method to construct the corresponding range of main color categories (the 11 color terms).³⁹ Overall speaking, studies from linguistics tradition tend to use color chips with unspecified light source as test materials. Studies in color science domain are more sophisticated in color specifications, but tend to overlook the psychological/cognitive constraints on the basic color terms. The experimental settings used by Shinoda *et al.* provide a reasonable bridge to link these two camps. Their results, while obtained from limited number of observers, also serve as a reference to which the color categorization behavior of other language speakers can be compared.

The purpose of this study is to demarcate the chromaticity range of the classic 11 color terms in Mandarin speakers. Moreover, the results would be compared with that of Shinoda's Japanese data. A sorting experiment was executed to specify basic color terms in the International Commission on Illumination (CIE) 1931 x - y chromaticity diagram (same as Shinoda *et al.*). Stimuli were square-sampled that vary across six different luminance (L) planes and vary in three perceptual dimensions: lightness, saturation, and hue. The effect of stimuli luminance and purity, which are related to perceptual dimensions of lightness and saturation, are examined for their potential interactions with color naming and categorization.^{10,11,39} The observers are all native Mandarin speakers who use traditional complex Chinese characters. The boundary between categories and each focal color,¹⁴ or centroid color,^{26,37} is the most representative exemplar within a color category and is demarcated and compared with those in studies of similar experimental design. The dependent variables are sorting items and response time (RT), which are submitted to serve indexes of central tendency (mode) and task difficulty, respectively. Similar measurements are sometimes collapsed to be unitary indexes such as “codability” or

TABLE I. Rank list derived from the survey of frequently used Chinese color terms.

Rank	Frequency count	Chinese color term	Phonetic transcription	English translation	Belonged color category (Berlin and Kay)	Denotation
1	128	紅	Hong	Red	Red	R
2	124	藍	Lan	Blue	Blue	B
3	106	紫	Zi	Purple/violet	Purple	P
4	95	粉紅	Fen-Hong	Pink	Pink	Pk
5	94	綠	Lu	Green	Green	G
6	90	橘	Ju	Tangerine	Orange	O
7	88	黃	Huang	Yellow	Yellow	Y
8	84	咖啡	Ka-Fei	Coffee	Brown	Br
9	69	桃紅	Tao-Hong	Deep pink	Pink	Dpk
10	65	灰	Hui	Gray	Gray	Gr
11	51	褐	He	Brown/tan	Brown	
12	48	棕	Tsong	Brown	Brown	
13	45	橙	Cheng	Orange	Gray	
14	44	青	Ching	Green/blue/black	Green/blue/black	
15	36	靛	Dian	Indigo	Blue	

The top 10 terms, plus black and white, were selected as semantic labels in the color sorting experiment.

“nameability,” which typically represent the observers’ consensus and dispersion.^{9,10,17,41}

The Mandarin translation of each basic color term is actually a thorny problem in this study, due to wide variance in Mandarin color terms across region, time, and speaker. In Berlin and Kay’s early survey on the development of color terms in worldwide languages,¹³ Mandarin has only four chromatic color terms in English: red, green, yellow, and blue. However, a recent study refutes this finding and provides evidences of a greater variety of color terms in use.¹⁸ There is currently no consensus on the conventional usage of color terms in contemporary Mandarin. Therefore, it is inappropriate to assign basic color names simply by following an existing dictionary. Additionally, the definitions of color terms in Chinese are more ambiguous than in English. There are diverse wording choices to describe the same color category. For example, the brown category can be conveyed by distinct terms like 咖啡 “Ka-fei” (the phonic translation of coffee), 棕 “Tsong” (palm fiber, coir), or 褐 “He” (tan). Similarly, multiple color categories can be expressed with identical color terms. The ancient polysemous term 青 “Ching” can refer to blue, green, and sometimes black. Although there are some studies concerning the usage of basic Mandarin color terms,^{17,18,36} the translations were unfortunately not consistent, particularly for nonlandmark basic terms like those excluding red, green, yellow, and blue. The term pink can be translated in two different ways: 粉紅 “Fen-Hong”¹⁷ and 桃 “Tao.”¹⁸ Brown can be both 棕 “Tsong”¹⁷ and 褐 “He,”¹⁸ and orange can be 橘 Ju¹⁷ and 橙 Cheng.¹⁸ To address this problem, the study first conducted a free-recall survey for filtering the current conventional Mandarin color terms. Only the terms that emerged most frequently from the empirical pretest were used as options in the sorting experiment.

METHOD

Experimental Design

A color sorting experiment using stimuli at six luminance (L) levels—5, 10, 25, 50, 100, and 170 cd/m^2 —is conducted to specify the range of major color categories on the CIE1931 chromaticity diagram. The sorting work is the force-choice task. Participants view the color stimulus under controlled viewing conditions, and then sort it into one of the 12 given color categories. In addition to the sorting items, RT is also recorded, because it reveals meaningful information about task difficulty.

There are 12 color categories to choose from, and these are labeled with traditional Chinese characters, that is, the original complex form instead of the Chinese simplified character. These categories were adopted as options for the color sorting due to their universality found in the landmark study¹³, and each corresponding color terms were carefully determined by an empirical survey of frequent Mandarin color terms. The traditional characters and phonetic transcriptions of these color terms are 紅 “Hong” (red), 橘 “Ju” (orange), 黃 “Huang” (yellow), 綠 “Lu” (green), 藍 “Lan” (blue), 紫 “Zi” (purple), 粉紅 “Fen-Hong” (light pink), 桃紅 “Tao-Hong” (dark pink), 咖啡 “Ka-Fei” (brown), 灰 “Hui” (gray), 白 “Bai” (white), and 黑 “Hei” (black). As discussed earlier, both synonymous and polysemous color terms are common in Mandarin. Moreover, the idioms of Mandarin color vocabulary vary across region, time, and speaker. Thus, the color terms used in labeling the 12 color categories were determined by a pretest on current popular color vocabulary rather than by arbitrary assignment. A ranking of Mandarin color term frequency counts, as listed in Table I, was obtained by having 133 participants perform a color vocabulary free-recall task. The voluntary participants are native Mandarin speakers aged 18–39 years, opportunity sampled from undergraduates, postgraduates, engineers,

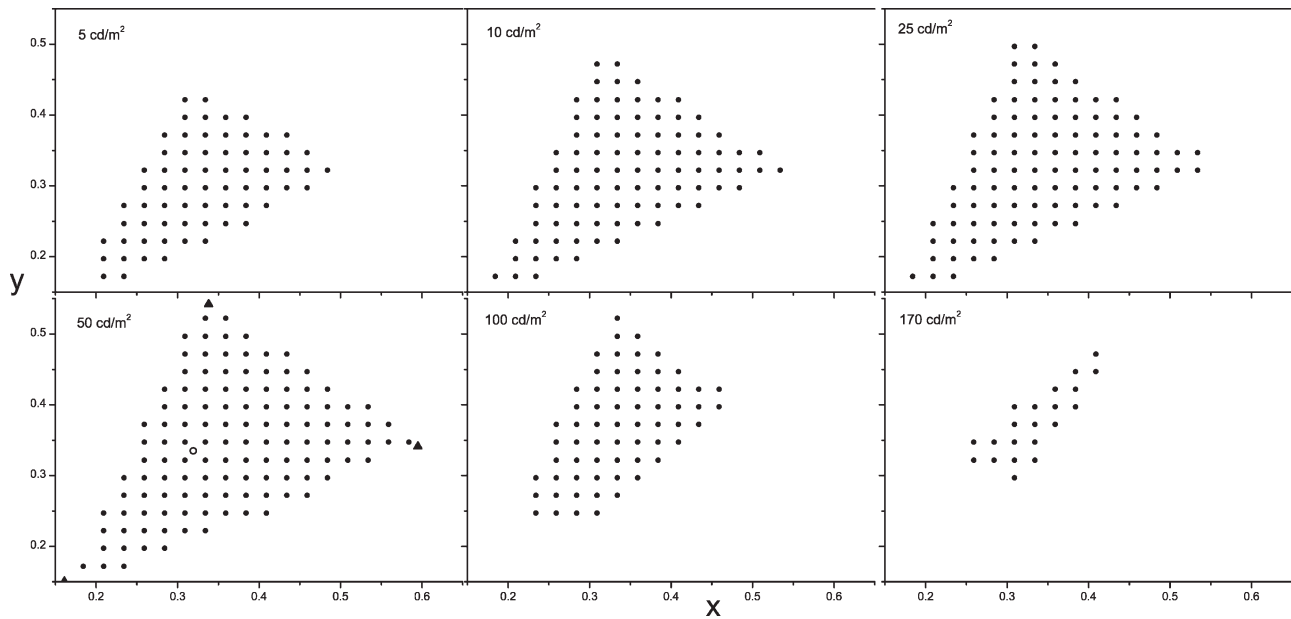


FIG. 1. Stimuli plotted on the CIE1931 L - x - y color space. Six sets of stimuli in distinct L levels were evenly spaced on x - y surfaces. The widest color gamut was measured at 50 cd/m^2 , and this gamut range is denoted with solid triangles. The center open circle in the 50 cd/m^2 plane represents the peak white (reference white) of the LCD monitor.

designers, home makers, school staff members, and teachers. The participants were instructed to “recall and write down monolexemic color terms you often use, hear and read.” This task was executed without any reference resources to elicit the most intuitive and tangible color terms currently in use. As shown in Table I, the top 15 terms comprise mostly universal hues,¹³ although some overlapping terms in similar categories are evident. The top 10 most frequently used Chinese color terms, corresponding to red, orange, yellow, green, blue, purple, pink, brown, and gray were selected as semantic labels in the sorting color task. Two terms, 粉紅 and 桃紅, which correspond to the English color term “pink” in related studies,^{17,18,36,42} are included because the authors assume that the two terms actually denote two distinct color categories according to cultural convention. This argument can be examined by the present color sorting experiment. In the pretest survey, black (黑 Hei) and white (白 Bai) were seldom counted as “color” terms, but still they were adopted as options in the sorting experiment. Consequently, a total of 12 Mandarin color terms serve as semantic labels for representing basic color categories; these are denoted in the study as R, O, Y, G, B, P, pk, dpk, Br, Gr, W, and Bk.

Participants

Forty-four participants (some of whom also participated in the pretest) screened with the Ishihara color vision test took part in the experiment. All are native Mandarin speakers aged 20–34 years, with 25 females and 19 males. Participants are undergraduate or postgraduate students at Chiao-Tung University, and their participation satisfied a course requirement. Participants have no formal

training in color science, and were not aware of the purpose or methodology of the study.

Stimuli

Six sets of colors were generated, corresponding to six different L levels: 5, 10, 25, 50, 100, and 170 cd/m^2 . Stimuli of the same L plane were evenly sampled along x - and y -axes in the CIE1931 x - y diagram. At each L plane, the sampling interval is 0.025 units, sweeping along the x - and y -axes to produce a regular and equal sampling of points within the gamut of display media. Six stimulus sets contain unequal amounts of colors—67, 89, 99, 121, 64, and 21, respectively—and these amounts depend on the availability of liquid crystal display (LCD) colors at different L levels. There are 461 distinct stimuli in total; all are plotted in Fig. 1. The widest color gamut constrained by the display media was measured at a level of approximately 50 cd/m^2 and is denoted with three solid triangles in Fig. 1. The open circle in the center of the 50 cd/m^2 plane presents the peak white ($x = 0.319$, $y = 0.335$, defined as the reference white in the study) of the display.

A look-up table was generated by the standard measuring procedure and the Matlab interpolation function was used to produce all color stimuli. The mean errors of all stimuli in chromaticity were also checked by the spectroradiometer; these were 4.42 cd/m^2 in L , 0.006 in x and 0.004 in y . The mean L error increases with the L level of stimuli. The mean errors of the stimulus sets of 5 and 10 cd/m^2 are 0.38 and 0.76, respectively.

The stimulus was displayed in a square, sized 2×2 visual angle at the required viewing distance of 50 cm. The background of the stimulus square was set at 60 cd/m^2 , the average L of all stimuli. In addition to the

TABLE II. The descriptive overview of categorical sorting results in different L level conditions.

	R	O	Y	G	B	P	Pk	Dpk	Br	Gr	W	Bk
5 (cd/m ²)	88	23	12	671	336	578	9	19	805	104	1	304
10 (cd/m ²)	203	47	34	1089	367	791	34	43	982	235	1	90
25 (cd/m ²)	348	437	39	1143	399	794	38	140	798	211	0	9
50 (cd/m ²)	248	786	265	1459	464	680	297	415	497	213	0	0
100 (cd/m ²)	15	284	331	1077	329	224	301	17	101	127	9	1
170 (cd/m ²)	4	20	300	255	124	23	87	2	22	39	46	2
Sum	906	1597	981	5694	2019	3090	766	634	3205	929	57	406

stimulus square, there was a thin black border surrounding the square and a white thin border surrounding the black border. Both the inner black and outer white borders were 0.025 wide, which are too thin to cause noticeable contrast effect. The setting of encircling stimulus with black or white borders, even other “decorative” color is common in many studies involving the assessment of color appearance, helps to produce a display with brightness or hue reference.¹⁰ The double border design in the experiment also intends to provide a layout with peak white and lowest L level of the display medium. Additionally, the border reduces the potential L contrast effect between the stimulus and the background, and holds the viewer’s attention to the stimuli.

Apparatus

Stimuli were presented and controlled by an ASUS F6E 13.3-inch laptop. Each stimulus was displayed in the exact center of the monitor. A well-calibrated PhotoResearch™ PR-650 SpectraScan spectroradiometer was used to measure all stimuli and the display characteristics of the LCD. The output uniformity stability check of the LCD was carried out according to a standardized procedure of 20 repeated measures. The measuring distance was 355 mm and the sample size was 10 cm² located in the center of the screen, and it covers the whole field of the spectroradiometer lens. The measuring geometry followed the recommendations of Photo Research, Inc. (Chatsworth, CA). The adopted standard colorimetric observer was CIE1931 and the reference white selected was D65. The mean maximum intensity at the center of the screen was 235.8 cd/m² [STD = 3.28, maximum value = 241 cd/m² (+2.11%), with a minimum value of 229 cd/m² (-2.96%)]. The mean maximum R , G , and B intensity and x - y value of the screen are: mean R = 52 cd/m², STD = 1.03, (x, y) = (0.595, 0.341); G = 140 cd/m², STD = 1.58, (x, y) = (0.338, 0.542); B = 44.6 cd/m², STD = 0.72, (x, y) = (0.161, 0.15).

Procedure

The experiment was conducted in a darkened room, with the only light source from the LCD. The viewing distance was set at 50 cm. The viewing distance and position were kept constant by an adjustable chin rest table with head fixer. There were a total of 461 trials. Participants were instructed to sort each stimulus into one of the given color categories. A custom keyboard with 12 tags of Mandarin

color vocabulary was used to input the sorting results. The initial 40 trials were for practice and were not recorded. The practiced observers were familiar with the position of color terms on the keyboard and were able to produce rapid and accurate sorting actions. All stimuli were presented in random succession. During the experiment, the observers could use a pause key and a resume key to break and restart the experiment. The flow of the experiment was controlled by Presentation® (Neurobehavioral System).

RESULTS AND DISCUSSION

Zone Map of Color Categories

The participants have produced 20,284 color category judgments via the force-choice sorting task. These judgments are submitted to render color zone maps connecting semantics with perception. Table II provides a descriptive overview of categorical sorting results in conditions of different L levels. In this table, the rank order for the sum of frequency counts is G (green), Br (brown), P (purple), B (blue), O (orange), Y (yellow), Gr (gray), R (red), Pk (pink), Dpk (deep pink), Bk (black), and W (white). There were very few unreasonable judgments—precisely two votes for Bk in 170 and one in 100 cd/m² conditions, and one for W in 5 and 10 cd/m². These should be ignored,

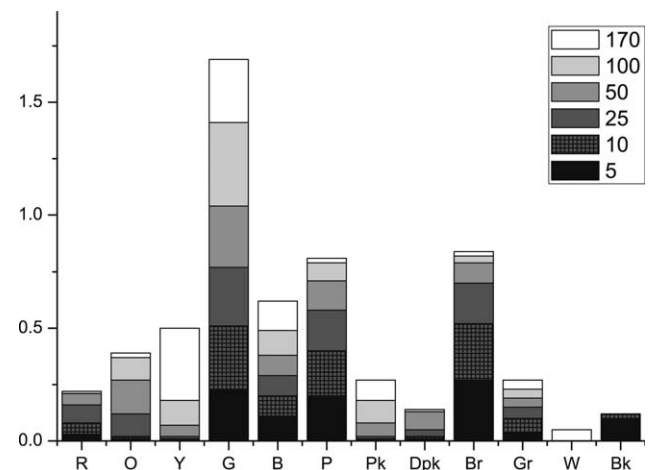


FIG. 2. The stack histogram for presenting normalized frequency distribution of each condition. The x-axis shows the color categories, and the y-axis shows the (accumulated) ratio. The various filled gray levels are used to denote six luminance conditions.

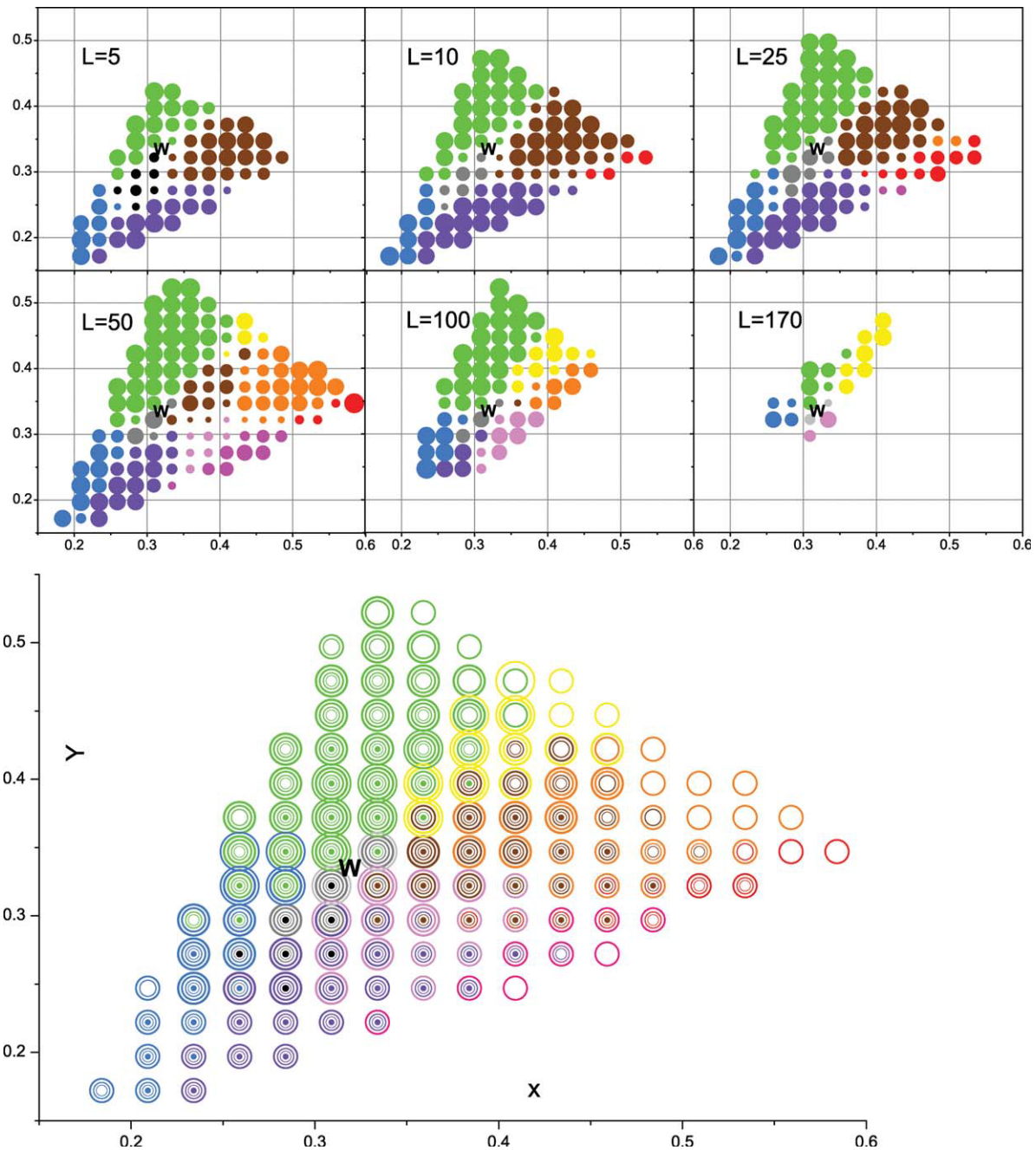


FIG. 3. Upper six x - y diagrams of different L levels using circle color and size to present color category and mode size, respectively. The lower diagram combines all results and differentiates the modes of six conditions with open circles that decrease in size. The larger outer circle represents the mode of the 170 cd/m^2 condition, and the smallest circle represents the mode of the 5 cd/m^2 condition.

because they may easily have been keyboard input errors. The input key for black was close to the key for white.

A cross- L comparison of the normalized frequency distribution is shown in the stack histogram in Fig. 2. The x -axis lists the given color categories, whereas y -axis presents the ratio of original counts to stimulus numbers of each condition, with various gray-level fills to differentiate the six L levels. The histogram presents a rough structure of the frequency distribution across color categories and L levels. The green, blue, and gray categories give relatively even frequency ratios, implying that these

three color concepts are luminance invariant; that is, they exist in all L conditions. In contrast, the other color categories are perceived at a limited number of L levels. Red, purple, and brown are more frequently perceived in medium to low L conditions, orange and deep pink are recognized in middle L conditions and yellow and pink are apparent in high L conditions. It is particularly noteworthy that the chosen color terms (except for the achromatic terms) are often referred to as “hue” terms, suggesting that these should be more or less independent from luminance and saturation. However, some “hue” terms, such

as yellow and deep pink, seems to be typical of specific luminance levels. This luminance-dependent phenomenon in color category sorting has been addressed³⁹ and will be discussed further in the following section.

The interaction between the L condition and the recognized color category is illustrated in Fig. 3. The upper six x - y chromaticity diagrams with colored circles represent the demarcated color zones in six L conditions. The coordinates of each circle's center correspond to those of the stimulus. The colors of the circles intuitively symbolize the category that gained maximum votes, the mode, except for the light gray in the $L = 170$ condition symbolizes white. In addition, the circle sizes correspond to the maximum number of judgments to visualize the degree of consensus better under each stimulus condition. Generally, larger circles symbolize the focal color of the category spread over the peripheral districts of high purity in colorimetry, whereas the smaller circles are found in the common border between distinct color zones and in the central area of low purity surrounding the reference white.

The composition of color zones appears to be diverse across the six L conditions. In the lowest L condition (5 cd/m^2), there are only five perceptually dominant categories: green, blue, brown, purple, and black. As luminance levels increase, the other color categories gradually become apparent. Specifically, red and gray become recognizable from $L = 10$, orange and deep pink from $L = 25$ and pink and yellow from $L = 50$. Certain color categories become less apparent in higher L conditions; specifically, red, deep pink, and brown are seldom identified from $L = 100$. In the highest L condition (170 cd/m^2), only yellow, pink, blue, green, and white remain visible. These results suggest that a common concept of color, labeled with a specific color term, is not merely an idea of a hue independent from brightness and saturation information. The results suggest that some color (hue) terms, such as red, yellow, pink, and others are strongly associated with luminance.

The lower diagram in Fig. 3 combines the above six diagrams. It reveals the spatial changes of color category compositions along with L level conditions. The color of the open circle represents the mode of color sorting and the size of the circle corresponds to the L level. The largest outer circles denote color zones designated in $L = 170$, and the inner smaller circles that decrease in size denote gradually decreasing L conditions. The color consistency of the concentric circles is an index of the degree of luminance dependency in the sorting results. Stimuli located around the borders between categories appear more ambiguous, and naturally are designated into different categories when luminance changes. The top and lower-left corner of the gamut triangle, demarcated as green, blue, and partial purple, show strong consistency across all L levels. However, color zones in the area from the center to lower-right corner of the triangle are strongly influenced by luminance conditions. It is important to note the superseding pattern of some groups in that area, such as a warm color group (yellow, orange, and

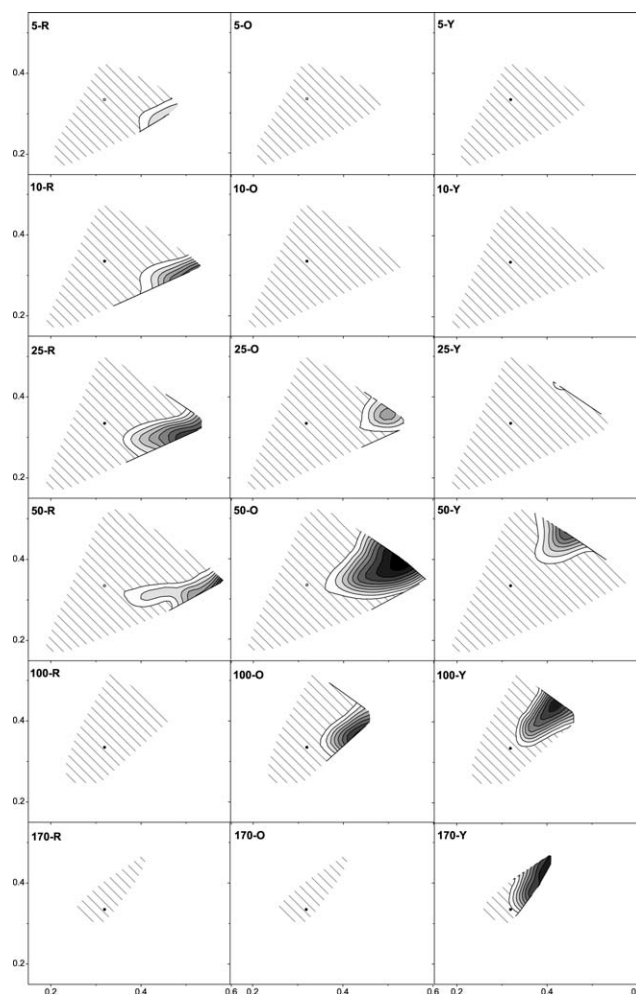


FIG. 4. Contour line map showing the formation of red, orange, and yellow. A detailed description is in the text.

brown) and another group (pink, deep pink, and purple). Members in these two groups seem to displace each other as the luminance conditions change. It is remarkable that the brown category overlaps a large range of color categories according to the fluctuation of luminance. A stimulus fixed in a chromaticity coordinate is recognized as brown in lower L but would be called yellow, orange or even pink as the L gets higher. In addition, the gray category also demonstrates a similar but weaker effect; it can substitute for many other colors as the luminance condition changes. This effect is related to the so-called “wild-card” phenomenon.⁴³

The particularized formation within each color category's luminance condition is presented in Figs. 4–7. Each figure contains 6×3 (the number of L levels by the number of color categories) diagrams of smoothed-out contour maps and all diagrams share the equivalent x - y unit and scale. The contour-smoothing algorithm was provided by OriginPro by OriginLab. The denotation of the conditions (cd/m^2 —color category name) is shown in top-left corner of each diagram. The small black dot in the middle of diagram marks the position of reference white. Different fills of gray level between contour lines repre-

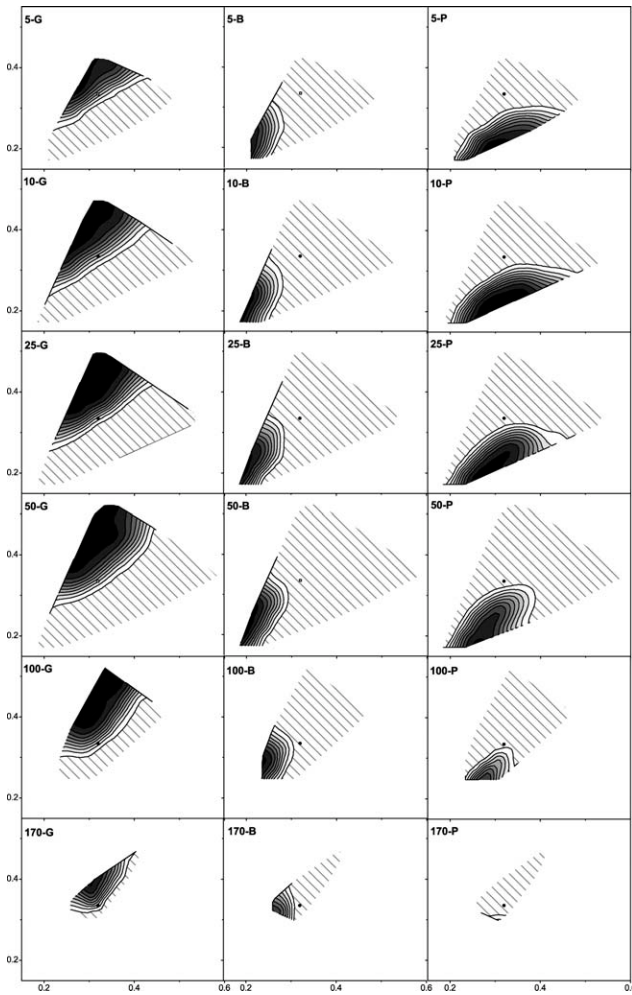


FIG. 5. Contour line map showing the formation of green, blue, and purple. A detailed description is in the text.

sent the frequency ratio of the observers' judgments. Zones filled with black indicate that these obtained over 90% of the votes for the corresponding color terms, which means they are focal zones with the least controversy. The other gray fills gradually increase in lightness according to their respective percentage of votes, with a decreasing interval of 10%. Consequently, the darker the zone fill, the higher the frequency, and the more representative the stimulus. The darkest gray fill indicates 80–90% votes, whereas the lightest fill (white) indicates 10–20% votes. A frequency ratio below 10% is ignored and filled with slash lines to mark the gamut. These contour line maps reflect the noticeable transformation of each color zone involving luminance variation. Each color category shows distinct pattern of emergence, congregation and lapse on the color space.

Figure 4 presents zones of red, orange, and yellow categories that show prominent lumiance-dependent features in their distribution. The red is recognizable below 50 cd/m^2 , and its focal zone (the zone with highest ratio) is relatively small, reaching only 70–80% ratio level. The covered area completely overlaps with the brown zone in 5 cd/m^2 , although the probability of seeing red at this L

level is quite low. The orange zone is also designated at restricted luminance levels, mainly in 50 and 100 cd/m^2 . The formation of the orange contour map reveals a very concentrated pattern; it contains a recognizable focal zone of over 90% in 50 cd/m^2 level, and then the zone diminishes drastically in 100 and 25 cd/m^2 levels. The range of orange and brown categories also overlap considerably, and brown also overlaps with yellow. The yellow zone is recognizable in conditions above 50 cd/m^2 , and the focal zone of over 90% can be found in 100 and 170 cd/m^2 level. The first three color categories discussed thus far contain colors of long wavelength range. Their territories are all luminance-dependent, and overlap with the brown zone in low luminance conditions. This result is consistent with the familiar perception that so-called warm colors (red, orange, and yellow) would shift into brown as they become darker.

Another cluster of color category zones—green, blue, and purple—is shown in Fig. 5. Apparently, viewers were able to perceive these three colors across all six luminance levels, except that purple was infrequently indentified in 170 cd/m^2 condition. The green category might be considered as a unique color concept that is particularly

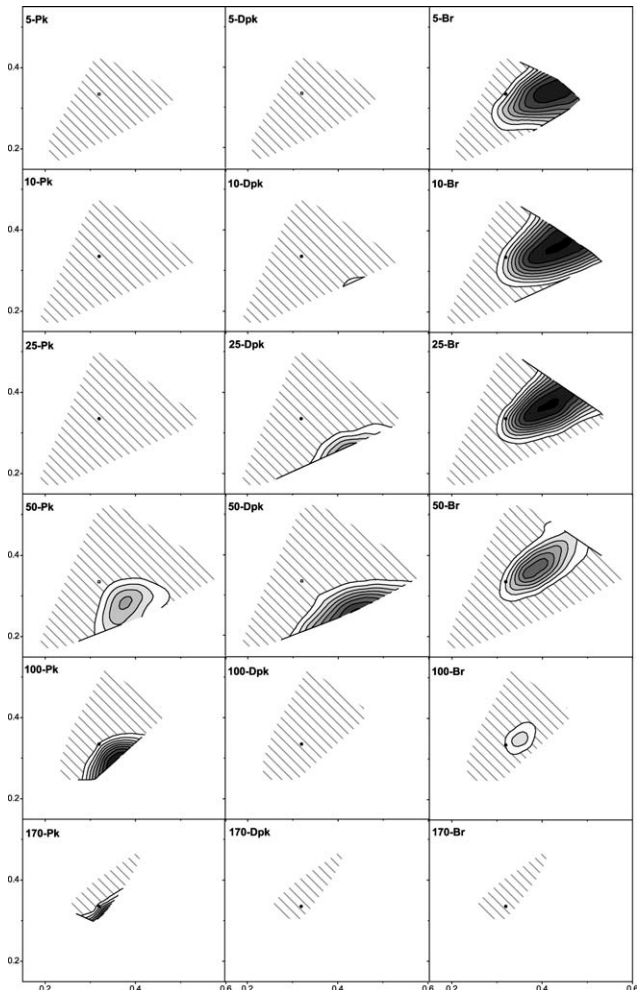


FIG. 6. Contour line map showing the formation of pink, deep pink, and brown. A detailed description is in the text.

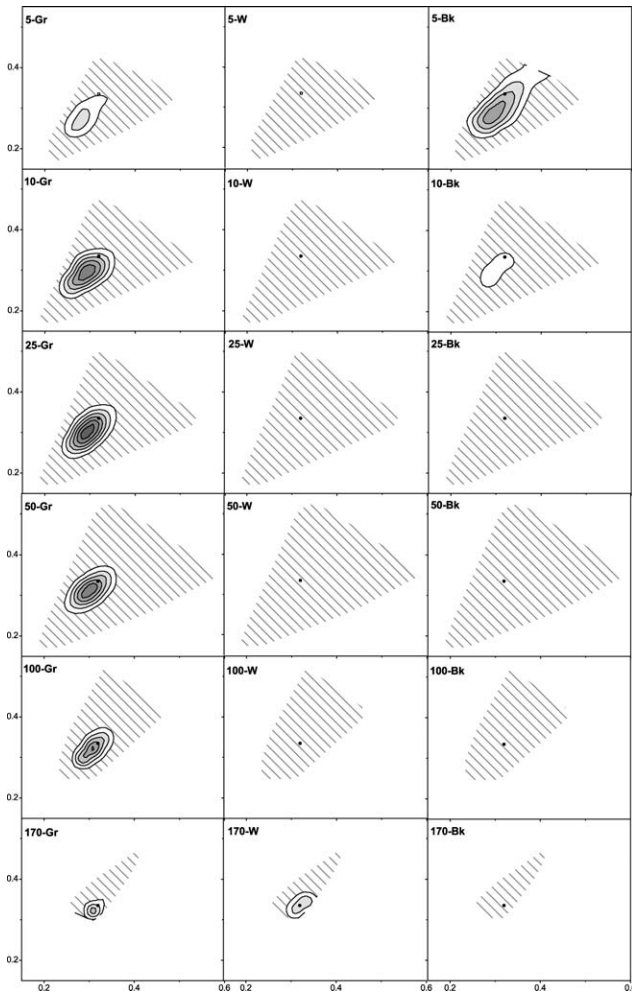


FIG. 7. Contour line map showing the formation of gray, white, and black. A detailed description is in the text.

easy to define, given its large focal zone of over 90% votes and sharp border. Its dense peripheral contour lines reflect a steep fall in frequency ratio. The location of the focal zone remains constant, rather than shifting with luminance changes. Moreover, the overlapping area between the green zone and neighboring purple, green, and brown zones is very limited in size. A similar pattern of contour lines can be found in the blue zone, although its covered range is much narrower than that of the green zone. The purple category is also a easy-to-identify color, as revealed in the concentrated pattern in the map, typically in conditions below 50 cd/m^2 . However, the perceptual definition of purple seems not as distinct as that of blue or green. In the darker conditions, its zone overlaps partially with those of brown and red, whereas in lighter conditions it overlaps with deep pink and pink. Generally, when compared with the previously discussed warm color cluster and the other color categories, the three colors in Fig. 5 demonstrate the notable characteristic of being perceptible across every L level. Furthermore, the overlapping zone between these and neighboring colors is relatively small, especially for the green and blue zones. All of these observed features suggest that the psychological

quality of these colors is more universal, stable and less ambiguous when compared with other colors in the study.

Figure 7 shows the contour map of the achromatic categories of gray, white, and black. The gray zone distributes around the lower-left area in all luminance situations, close to the intersection of the blue, purple, green, and brown zones. It also appears more clearly in the middle luminance levels, and switches to black in the lowest luminance conditions. The term of white was used only in 170 cd/m^2 condition, and its zone encircles the white point. White and black categories gained very low votes, perhaps due to the fact that observers were provided with thin outlines of white and black for reference with each stimulus. Literally, gray and black should be neutral color concepts that do not involve any hue information. However, the results show gray as a category that represents the “cold” cluster of colors, typically blue and purple, in very low saturation conditions, and black corresponds to cold colors in very low saturation and luminance conditions. The actual neutral exemplar in any luminance level should be located around the reference white point, just as in the white zone. Based on the present results, the ideal neutral point lies on the border between the brown and gray/black zones.

Response time and Boundary Definition

With the constrained option of 12 color terms, the perceptual regions corresponding to the main color categories on the CIE x - y diagram were carefully mapped out, as shown in all previous figures. However, the regions defined in those figures, such as the distinct color zones seen in Fig. 3, were based on one single statistic: namely, the quantity of votes for a certain color term. Another way to help define the boundary between color territories is to take into account the task difficulty measure. In this study, the RTs in each sorting trial are rendered as dependent factors relative to the ease in making a color category judgment. It is assumed that the more ambiguous the color, the longer it takes to discriminate and sort the color into one of the given categories. The RTs were also considered to reflect the inner categorical structure in the related studies.^{9,10,44} While the size of the mode is an index of the commonness of the stimulus, the RT is an index of the distinctiveness of the stimulus. A stimulus that results in a rapid response plus a larger mode to the same color term signifies that it is well located in the center zone of a color category (i.e., it is a typical example of that category). The reverse situation, with a long RT and fewer votes, indicates a stimulus located in the periphery of a category or the boundary between categories.

Figures 8 and 9 visualize two factors: the 50%/75% vote threshold and the contour map of RTs, respectively. Both figures contain six luminance levels in the x - y diagram of the same scale. Fig. 8 uses a unitary criterion to demarcate the boundary of color zones: namely, the vote frequency counts of 75% level (color fills) and 50% level (color lines). Figure 9 presents RT in terms of contour lines on the color space. The black area corresponds to

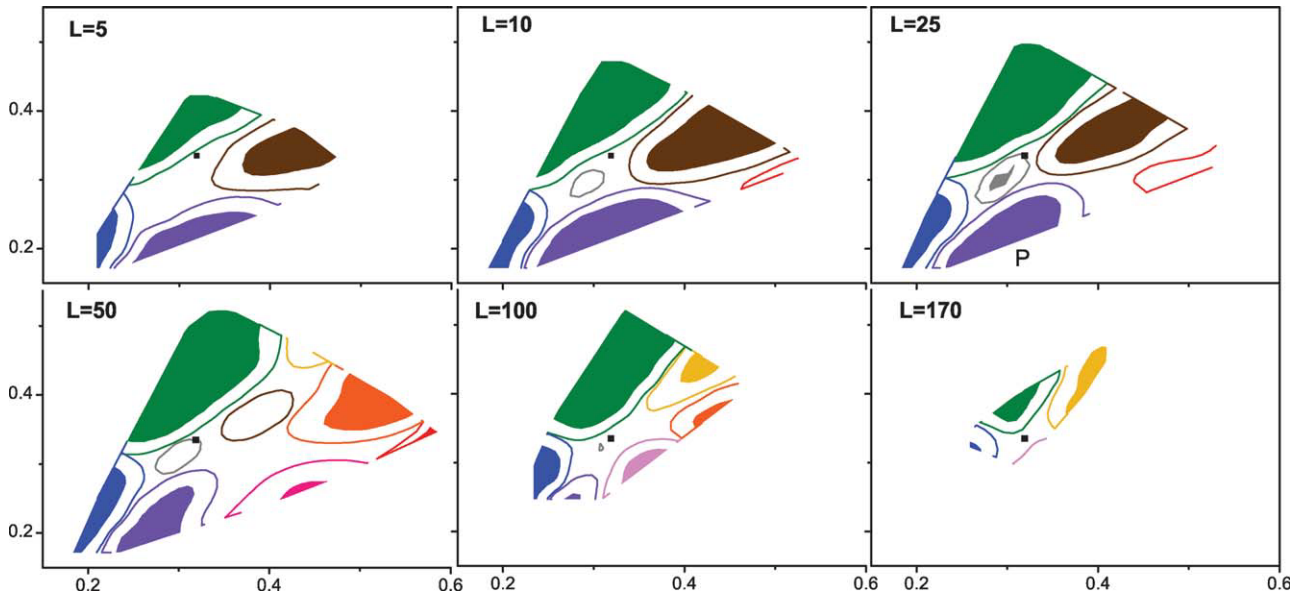


FIG. 8. Zones of color categories in six luminance conditions. The boundaries are demarcated by 75% and 50% votes ratio, which are marked with color fills and color lines, respectively. These two threshold levels partition the x - y surface into distinct zones without overlapping.

RT below 1.5 s, whereas the white area corresponds to time beyond 2.25 s.

It is interesting to examine the connection between the spatial constitutions in these two figures. The center tendency index can demarcate the core zone of the category, as shown in Fig. 8, whereas the RTs information gives robust weight to the boundary. In Fig. 9, there are several prominent hot spots (black areas) embedded in the inert ground (white areas). The white and lightest gray areas, representing long RTs, are generally consistent with the areas of achromatic center and boundaries between cate-

gories in Fig. 8. The Pearson correlation coefficient of the size of mode and the RT of each stimulus in low to high L conditions is -0.846 ($P < 0.01$), -0.874 ($P < 0.01$), -0.816 ($P < 0.01$), -0.81 ($P < 0.01$), -0.763 ($P < 0.01$), and -0.103 ($P = 0.67$), respectively. High and stable negative correlation can be observed in most conditions. This effect disappears only in the $L = 170$ condition. Generally, the relation between measured RTs and color boundaries is consistent with Guest's study adopting RTs as one of the quantified index presenting the perceptual boundaries in the free color-naming behavior.¹⁰

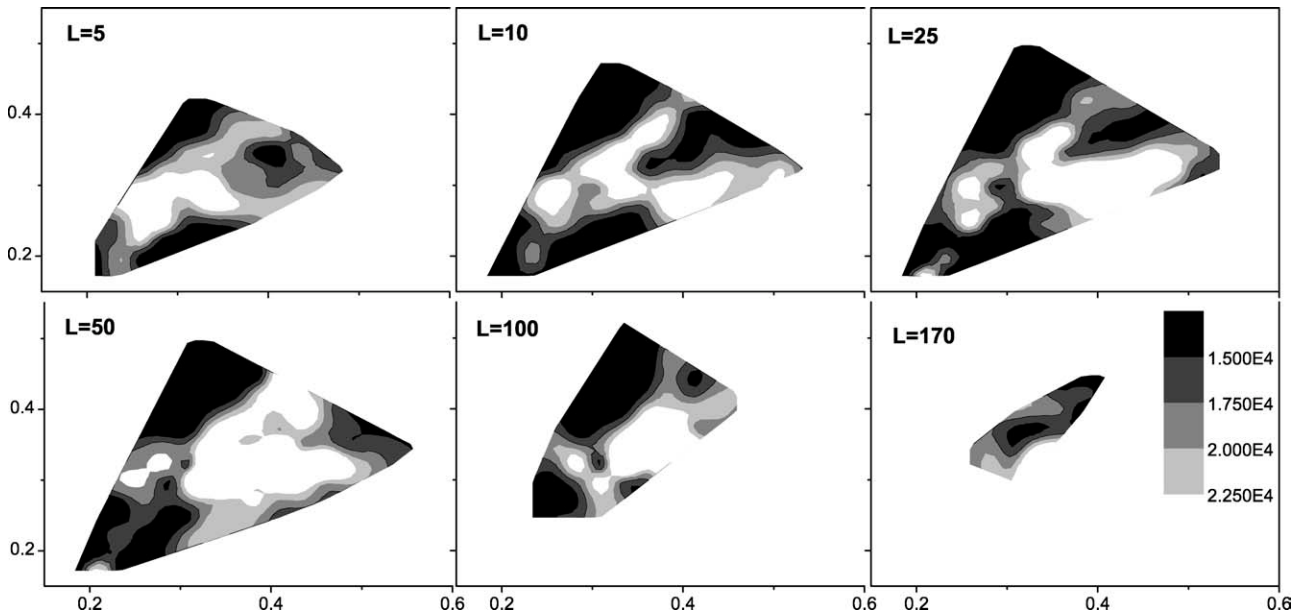


FIG. 9. Contour maps presenting RTs in six luminance conditions. The light areas indicate longer RTs, or the more difficult sorting decisions, whereas black areas indicate the faster RTs and easier response zones. The darker areas roughly correspond to the color zones in Fig. 8, except in $L = 170$ condition.

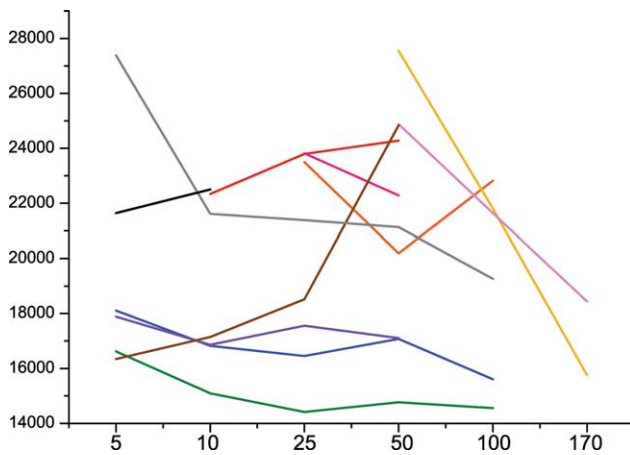


FIG. 10. Line plot of the mean RTs of 11 color categories (white is excluded) in their frequently identified luminance conditions. The y-axis shows milliseconds and the x-axis shows luminance.

The RT measure also reveals the distribution of perceptual distinctiveness (saliency) on the color space. Figure 10 presents the luminance-against-RT line plot that connects the mean RT of the color categories in certain L conditions. Note that the figure does not contain every category in every condition. To prevent the interference of the RT of non-typical judgments, each line of color category only presents the L conditions in which obtained votes surpass 10% of all votes within the category. The white category is not included, because its votes ratio reaches 10% only in $L = 170$ condition. The line plot shows that the RT is both category- and luminance-related. For the categories of green, blue, and purple, the mean RTs are generally shorter across all L conditions. This suggests that observers can easily and rapidly decide whether a given color belongs to the green, blue or purple categories, even though these three are next-door neighbors on color category maps. However, the lengths of RT in the other categories vary with luminance. The RTs of gray drop drastically, indicating that gray is easier to determine in higher L levels, whereas RTs of brown show the reverse trend. The rest of the color categories also have shorter RTs in their corresponding dominant L levels. The mean RT of the orange category, for instance, drops at $L = 50$, which is the luminance level at which the color is most frequently recognized.

Rapid RTs can be found in the highest L condition of $L = 170$, as shown in Figs. 9 and 10. This appears to be somewhat in conflict with the previous point that RT serves as an index of the ease level of the task, as a color displayed in very high luminance should decrease in the perception of saturation,⁴⁵ and it should thus become more difficult to determine its appropriate hue category. However, there are two possible reasons for the actual result. First, the limitation of the display gamut makes the colors of high L conditions vary in restricted numbers of categories. The second reason is that under such high luminance conditions, the observers actually make a color-or-white distinction, that is, they simply sort the stimulus into one of two main categories. The psychologi-

cal distance between these two categories should be larger than that between many other color categories, such as green and yellow. With the limitation of the gamut display, the number of subcategories under the broader “color” category is even fewer, as designated by the yellow, green, blue, and pink zones in Fig. 8. These factors could reduce the task difficulty in $L = 170$ condition and contribute to the quick response.

SUMMARY

The study presents the formation of color categories through a 12-color-terms sorting experiment that uses native Mandarin speakers as participants. The adopted categorical color terms were determined to be universal among human cultures,^{10,13,16–18,36,39,42} and were confirmed to be frequently used by a free-recall pretest. The range of each term-related color category among observers was carefully plotted on the CIE 1931 chromaticity diagram on six luminance planes. Unlike many studies adopting reflective materials, limited saturation or luminance setting, or irregular sampling in designing stimuli, this study’s illuminant stimuli vary regularly in terms of hue, lightness, and saturation and can systematically capture the spatial structure of color categories in different perceptual dimensions. In general, this experimental design leads to an intriguing finding in the results: namely, the changing shape of the color zone depending on purity and luminance. These two colorimetric parameters correspond roughly to saturation and lightness. In the seminal *Color Categories in Thought and Language*,¹¹ Jameson and D’Andrade argue that within the internal perceptual color space, hue interacts with saturation and lightness to produce “bumps.” Bumps are defined as the salient representation of color categories or the foci colors. The formation of focal color zones located at different luminance levels and eccentricities (see Figs. 4–7) apparently support and “visualized” this theory.

The formation of color categories shows the various degrees of the luminance effect. The most luminance-irrelevant cluster includes green, blue, purple, and gray. These four colors, particularly green and blue, are identified across all luminance levels. Additionally, the shape of the corresponding contour map remains stable, and the location of the foci of these categories is consistent across all conditions. Moreover, the RTs of the green, blue, and purple categories are the shortest among all colors, and are unrelated to variances in luminance. All measures indicate that these three color concepts, particularly green, are more psychologically distinctive, salient and robust than others. Green gained the most votes in the experiment with the lowest mean RT, and its zones are encircled by sharp contour edges. Interestingly, the locations of these three categories on the color space are close. Blue is adjacent to green and purple is adjacent to blue. These color categories are similar in chromaticity, but distinct in category distinguishing. Nevertheless, many categories can be frequently identified and appear typical

only in certain restricted luminance ranges. Red is typical in $L = 10\text{--}25$, deep pink in $L = 50$, orange in $L = 50\text{--}100$, pink in $L = 100$, and yellow in $L = 100\text{--}170$. Conceptually, these color categories are different shades of the “warm” color cluster and are bound tightly by luminance conditions. In low luminance levels, the same chromaticity location of warm colors could easily be identified as brown. Additionally, the red, deep pink and pink categories, which belong to the “Hong” (red) cluster in Mandarin, appear to be typical in three distinct ascending luminance levels. Their foci locations do not overlap. These factors indicate that Hong, Fen-Hong, and Tao-Hong could be independent categories among Mandarin speakers. Also, the claims of earlier studies of Mandarin, which accounted only for six color categories,¹³ could be inappropriate to apply to the contemporary Mandarin environment. In Berlin and Kay’s survey on the development of color terms in worldwide languages, Mandarin has only four chromatic color terms: red, green, yellow, and blue. Some researchers argue that these limitations are refutable and have tried to propose new evidence.¹⁸

Furthermore, it is important to note that the foci of brown and gray are located symmetrical to the reference white. Traditionally, gray should serve as a representation of achromatic stimuli, but the results show that it actually stands in for “cold” colors in low saturation conditions, whereas brown stands in for warm colors in similar conditions. The exact neutral gray may only exist in perfectly controlled viewing conditions, which are seldom found in the real world. Supposedly, these two wild-card color concepts⁴³ are capable of conveying near achromatic shades of cold- and warm-tinted colors.

The present results are comparable with the study that uses similar viewing conditions and color space but employed Japanese native speakers’ as observers.³⁹ The most intrigued contrast in the comparison is that the distinct location of boundaries between blue and green are different than the location observed in these two studies. The green areas in this study’s color zone maps extended further than the blue areas, whereas the reverse was found in the comparative study. Also, the red area in this study is narrower than that in the Japanese study. Other than the differences in area size of red and the boundary location of blue and green, the remaining color categories were similarly spaced in both studies. Interestingly, blue and green can be loosely represented by a term in a literary language used by ancient Chinese, and this ancient Chinese written language influenced both modern Mandarin and Japanese. Perhaps, the conventional definitions of blue and green in modern Mandarin and Japanese developed differently. In conclusion, a greater quantity of substantial empirical data would undoubtedly improve the overall understanding of the categorical color-naming issue.

1. Bornstein MH. Perceptual categories in vision and audition. In: Har-nad S, editor. *Categorical Perception—The Groundwork of Cognition*. Princeton, New Jersey: Cambridge University; 1990.
2. Bornstein MH, Korda NO. Discrimination and matching within and between hues measured by reaction times: Some implications for

- categorical perception and levels of information processing. *Psychol Res* 1984;46:207–222.
3. Boynton RM, Fargo L, Olson CX, Smallman HS. Category effects in color memory. *Color Res Appl* 1989;14:229–234.
 4. Goldstein J, Davidoff J, Roberson D. Knowing color terms enhances recognition: Further evidence from English and Himba. *J Exp Child Psychol* 2009;102:219–238.
 5. Gouras P. Color vision. In: Kandel ER, Schwartz JH, Jessell TM, editors. *Principles of Neural Science*, 3rd edition. New York: Elsevier; 1991.
 6. Heider ER. Universals in color naming and memory. *J Exp Psychol* 1972;93:10–20.
 7. Uchikawa K, Shinoda H. Influence of basic color categories on color memory discrimination. *Color Res Appl* 1996;21:430–439.
 8. Seliger HH. Measurement of memory of color. *Color Res Appl* 2002;27:233–242.
 9. Sturges J, Whitfield TW. Salient features of Munsell colour space as a function of monolexic naming and response latencies. *Vision Res* 1997;37:307–313.
 10. Guest S, Van Laar D. The structure of colour naming space. *Vision Res* 2000;40:723–734.
 11. Jameson K, D’Andrade R. It’s not really red, green, yellow, blue: An inquiry into perpetual color space. In: Hardin C, Maffi L, editors. *Color Categories in Thought and Language*. Cambridge, UK: Cambridge Univ. Press; 1997.
 12. Regier T, Kay P, Gilbert A, Ivry R. Language and thought: Which side are you on, anyway? In: Malt B, Wolff P, editors. *Words and the Mind: Perspectives on the Language–Thought Interface*. New York: Oxford University Press; 2010.
 13. Berlin B, Kay P. *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press; 1969.
 14. Regier T, Kay P, Cook RS. Focal colors are universal after all. *Proc Natl Acad Sci USA* 2005;102:8386–8391.
 15. Kay P, Regier T. Color naming universals: the case of Berlinmo. *Cognition* 2007;102:289–298.
 16. Lindsey DT, Brown AM. Universality of color names. *Proc Natl Acad Sci USA* 2006;103:16608–16613.
 17. Lin H, Luo MR, MacDonald LW, Tarrant AWS. A cross-cultural colour-naming study. Part I. Using an unconstrained method. *Color Res Appl* 2001;26:40–60.
 18. LüChing-Fu. Basic Mandarin color terms. *Color Res Appl* 1997;22:4–10.
 19. Bernasek JL, Haude RH. Color-naming and color-matching by pre-school children as a function of visual field and sex. *Percept Mot Skills* 1993;77(3 Part 1):739–747.
 20. Bornstein MH. On the development of color naming in young children: Data and theory. *Brain Lang* 1985;26:72–93.
 21. Karpf RJ, Gross AE, Small MY. Naming, selection, and ordering of color (“hue”) by young children. *J Gen Psychol* 1974;90(2d Half):297–314.
 22. Pitchford NJ, Mullen KT. Is the acquisition of basic-colour terms in young children constrained? *Perception* 2002;31:1349–1370.
 23. Sakurai M, Ayama M, Kumagai T. Color appearance in the entire visual field: Color zone map based on the unique hue component. *J Opt Soc Am A Opt Image Sci Vis* 2003;20:1997–2009.
 24. Holmes A, Franklin A, Clifford A, Davies I. Neurophysiological evidence for categorical perception of color. *Brain Cogn* 2009;69:426–434.
 25. Boynton RM, Gordon J. Bezold–Brücke hue shift measured by color-naming technique. *J Opt Soc Am* 1965;55:76–78.
 26. Boynton RM, Maclauray RE, Uchikawa K. Centroids of color categories compared by two methods. *Color Res Appl* 1989;14:6–15.
 27. Ingling CR, Jr., Scheibner HM, Boynton RM. Color naming of small foveal fields. *Vision Res* 1970;10:501–511.
 28. Boynton RM, Schafer W, Neun ME. Hue–wavelength relation measured by color-naming method for three retinal locations. *Science* 1964;146:666–668.
 29. Boynton RM, Olson CX. Salience of chromatic basic color terms confirmed by three measures. *Vision Res* 1990;30:1311–1317.
 30. Roux FE, Lubrano V, Lauwers-Cances V, Mascott CR, Demonet JF. Category-specific cortical mapping: Color-naming areas. *J Neurosurg* 2006;104:27–37.

31. Dedrick D. The foundations of the universalist tradition in color-naming research (and their supposed refutation). *Philos Soc Sci* 1998;28:179–204.
32. Roberson D, Davies I, Davidoff J. Color categories are not universal: Replications and new evidence from a stone-age culture. *J Exp Psychol Gen* 2000;129:369–398.
33. Ozgen E, Davies I. Acquisition of categorical color perception: A perceptual learning approach to the linguistic relativity hypothesis. *J Exp Psychol*. 2002;131:477–493.
34. Kay P, Regier T. Language, thought and color: Recent developments. *Trends Cogn Sci* 2006;10:51–54.
35. Regier T, Kay P, Khetarpal N. Color naming reflects optimal partitions of color space. *Proc Natl Acad Sci USA* 2007;104:1436–1441.
36. Lin H, Luo MR, MacDonald LW, Tarrant AWS. A cross-cultural colour-naming study. Part II - Using a constrained method. *Color Res Appl* 2001;26:193–208.
37. Sturges J, Whitfield TWA. Locating basic colours in the munsell space. *Color Res Appl* 1995;20:364–376.
38. Grant DW. Visual asymmetry on a color-naming task: A developmental perspective. *Percept Mot Skills* 1980;50:475–480.
39. Shinoda H, Uchikawa K, Ikeda M. Categorized color space on CRT in the aperture and the surface color mode. *Color Res Appl* 1993;18:326–333.
40. Kay P, Berlin B, Maffi L, Merrifield W. Color naming across languages. In: Hardin C, Maffi L, editors. *Color Categories in Thought and Language*. New York: Cambridge Press; 1997.
41. Zollinger H. Categorical color perception: Influence of cultural factors on the differentiation of primary and derived basic color terms in color naming by Japanese children. *Vision Res* 1988;28:1379–1382.
42. Lin H, Luo MR, MacDonald LW, Tarrant AWS. A cross-cultural colour-naming study. Part III - Using colour-naming model. *Color Res Appl* 2001;26:270–277.
43. Greenfield PJ. What Is gray, brown, pink, and sometimes purple: The range of "Wild-Card" color terms. *Am Anthropol* 1986;88:908–916.
44. Guest S, Van Laar D. The effect of name category and discriminability on the search characteristics of colour sets. *Perception* 2002;31:445–461.
45. Capilla P, Luque MJ, Gomez J, Palomares A. On saturation and related parameters following guth's ATD colour-vision model. *Color Res Appl* 2001;26:305–321.