https://doi.org/10.2352/issn.2169-2629.2021.29.247

G_0 Revisited as Equally Bright Reference Boundary

Hao Xie and Mark D. Fairchild;
Program of Color Science / Munsell Color Science Laboratory, Rochester Institute of Technology, Rochester, New York USA

Abstract

Brilliance and zero grayness (denoted as G_0) and are two terms coined by Ralph Evans. Nayatani, Heckaman and Fairchild have done series of work to incorporate them into comprehensive color appearance models. In this work, those concepts were reexamined to scale lightness/brightness across the chromaticity diagram. Specifically, observers, mostly with a color science background, were asked to adjust the luminance of a color patch to appear with no grayness, or equivalently just about/cease to glow. The hypothesis was that lightness can be equalized across those chromaticities and the Helmholtz-Kohlrausch effect is automatically incorporated. This hypothesis was verified in a follow-up experiment where another group of observers completed paired comparisons of the brightness between the collected G_0 results. The G_0 task was also repeated under another two levels of adaption backgrounds, based on which different absolute brightness results for a given chromaticity might be derived. In addition, high correlations between the G_0 results (as a perceptual boundary between appearance modes) and different physical gamut boundaries including MacAdam's optimal colors were found for possible computational proxies and ecologically meaningful implications.

Introduction

(Chromatic) Brightness/Lightness Model

Brightness and lightness, defined as "attribute of a visual sensation according to which an area appears to emit more or less light" and "the brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white" [1], are of fundamental importance in color appearance modeling. From the definitions, lightness perception is based on brightness perception with a relative normalization (although the "white" is sometimes hard to define). Only related colors as opposed to unrelated/isolated colors possess both brightness and lightness. In photometry, brightness has been approximately represented by the physical metric, luminance, which is calculated by integrating spectral radiance with the CIE 1924 standard observer or its mathematical proxy $V(\lambda)$. And, the widely used lightness metric CIE L^* (in CIELAB and CIELUV) is a function of the luminances of the target stimulus and a (pre-)specified "white". $V(\lambda)$ has been a practical success for its additivity (Abney Law). However, it has been discussed that $V(\lambda)$ works well only for specific stimulus conditions where it was derived (i.e., high temporal and low spatial frequency in heterochromatic flicker photometry) [2, 3]. The results from a more natural psychophysics scheme, heterochromatic brightness matching (HBM), did not lead to a $V_b(\lambda)$ function with additivity property [4] and systematically consistent results across observers [5]. Also as phenomenologically noted in the Helmholtz-Kohlrausch (H-K) effect, the brightness to luminance ratio is not equal across the chromaticity gamut; instead, more saturated colors appear to be brighter with a hue dependency [1]. In other words, brightness, hue, and saturation are not independent within current color appearance models; the brightness of chromatic stimuli partly comes from their chromatic components, when their luminances are held constant to a neutral reference. To quantify the effect, equations have been proposed for unrelated colors (Ware & Cowan system) and related colors [6]. However, according to a recent review on brightness modeling, none of the state-of-the-art models provide satisfactory results for the H-K effect [7].

Brilliance and G_0 Functions

A different approach to modeling chromatic colors' brightness/lightness is to combine (in a colorimetric sense) brightness/lightness and saturation into a new perceptual attribute, "brilliance" coined by Evans. In his experiment [8], a series of monochromatic stimuli were centered with a neutral background, and the observer (Evans himself) adjusted the luminance of the center until it appeared in a mode between object color and self-luminous color, a state which Evans called "fluorence" (a perception to be distinguished from the physical "fluorescence" [9]) or equivalently color with zero gray content. Those threshold luminances for each wavelength are their G_0 functions. And such a concept can be generalized to the entire chromaticity diagram, as done for the H-K effect [5]. According to Fairchild [1], brilliance is some kind of apparent brightness that incorporates the H-K effect, and G_0 defines the luminance of "equal chromatic brightness (really, just brightness)" for various chromaticities.

Although no simple mathematical model of G_0 has been developed, its significance was further studied, especially by Nayatani and Heckaman & Fairchild [10, 11, 12, 13]. Speigle and Brainard have connected Evans' work with chromatic adaptation [14]. In their experiments, observers adjusted the luminances of different chromaticities under different ambient conditions until the stimuli appeared self-luminous. And by comparing the method of adjusting and yes-no staircase, no systematic differences had been found. Note that luminosity is slightly different than Evans' fluorence [9]. The boundary between reflective and self-luminous appearance modes is of great importance. In the movie industry, when colorists (over-)adjust the brightness/saturation of local objects, especially on a high-dynamic-range & wide-color-gamut display where brilliance is more relevant, the colors may appear fluorescent. Such warning boundary is approximately the G_0 gamut boundary [15, 16].

G_0 as Brightness Model Anchors

G_0 luminance only corresponds to a single point on a brightness scale. If extrapolating from Nayatani's hypothesis, all the G_0 luminances across the chromaticity diagram may appear equally bright. Then it would be interesting to consider G_0 as an anchor for a specific chromaticity, that is, instead of normalizing the luminance of the target stimulus to a white luminance, it may be more valid to do normalizations with respect to G_0 luminances. Fairchild and Heckaman have provided such framework and preliminary implementations [17, 18], where the G_0 luminances were found using NCS blackness, if not visually.

In this work, the concepts of brilliance and zero grayness were revisited by two psychophysical experiments. Specifically, in the first experiment, the G_0 luminance data for a group of representative chromaticities were collected under different adaptation backgrounds, which are significantly higher than previous related work for a potential high-dynamic-range relevant model. The results were compared across both chromaticities and adaptations, and against two physical luminance boundaries with high correlations. Whether, for a given adaptation level, those G_0 luminances appear equally bright was tested with a paired comparison experiment. The results show that the brightness scale is equalized, although not perfectly, as the differences for most pairs of G_0 colors are not statistically different. We describe the details of our experiments in the next section, followed by result analysis and discussion. In the last section, we summarize our main findings.

Experiments

General Settings and Stimulus Selection

Our work has received approval from our Institutional Review Board. Both experiments were conducted in a dark room (with most surfaces covered in black) where the stimuli were presented on an Apple Pro XDR display. The display has a peak luminance of $\sim$1500 cd/m^2 and a color gamut approximately DCI-P3. The graphical user interface was programmed in Apple's Swift language and the dynamic range and color gamut/encoding were managed by the Metal API (`https://developer.apple.com/metal/`). The characterization accuracy based on a linear model (3-by-3 matrix) and three 10-bit RGB look-up-tables (LUTs) [19] achieved an average of $\sim$0.57 ΔE_{00} for randomly sampled colors across the gamut, which was considered adequate for our objectives. Every time before the experiments, the display was warmed up and re-calibrated to match the characterization condition.

Figure 1 shows the stimulus configuration used in the first experiment. The center square patch covered a 3-by-3 deg field size and the rest of the full screen was filled with random neutrals as the background, corresponding to a field size of 39-by-22 deg. The observer was seated in front of the center of the display at a distance of 1 meter and adjusted the seat height to make their eye level to the stimuli level. The background's lightness levels ranged from L^* of 0 to 100 at a uniform interval of 25 so that the average was linearly integrated to L^* of 50 or $\sim$18% gray [20]. The absolute luminance levels were subject to the peak pixel in the background set for different adaptation levels, which will be described in more detail. The display's native resolution was 6016-by-3384 and the smallest unit of the random neutrals had 10 pixels. For the second experiment, two 3-by-3 deg color patches were similarly placed in the display center with a separation of one deg visual field between them. On the top of the screen, there was a text box showing the instruction. Particularly, for the first experiment, arrows were used to indict the adjustment options provided, which will be further explained in the next section.

Given the display gamut constraint, a group of chromaticities corresponding to representative colors from the Munsell color system were selected. For each hue (5R (red) / Y (yellow) / G (green) / B (blue) / P (purple)), the maximum chroma level within the display gamut was first determined, then along constant Munsell hue and lightness value three chroma levels were sampled. Thus the chromaticities of those 15 Munsell colors plus the perfectly reflective white under CIE D65, plotted in Fig. 2 and listed in Table 1, were used as the stimuli for finding their G_0 luminances. Other chromaticities might be interpolated or extrapolated from their results.

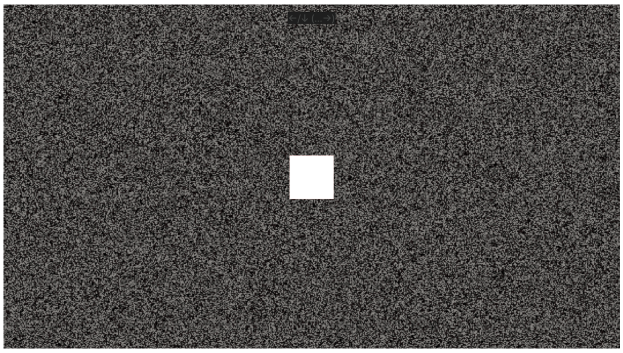

Figure 1. *Stimulus configuration in the first experiment. The center patch as the target was adjusted to the G_0 level. For the second experiment, two 3-by-3 deg color patches were similarly placed in the display center with a separation of one deg visual field between them. On the top of the screen, there was a text box showing the instruction (see details in the main text).*

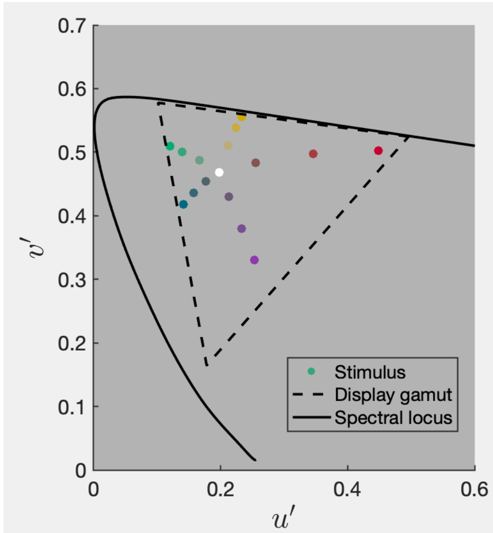

Figure 2. *Stimulus chromaticities versus display gamut.*

Table 1. Stimulus chromaticities and their Munsell specifications (when at particular relative luminances Y).

Stimulus	CIE 2-deg $u'v'Y$ under CIE D65	Munsell
1	(0.4480, 0.5024, 11.7)	5R 4/16
2	(0.3453, 0.4970, 11.7)	5R 4/10
3	(0.2548, 0.4829, 11.7)	5R 4/4
4	(0.2333, 0.5556, 42.0)	5Y 7/12
5	(0.2242, 0.5385, 42.0)	5Y 7/8
6	(0.2117, 0.5104, 42.0)	5Y 7/4
7	(0.1201, 0.5092, 29.3)	5G 6/11
8	(0.1400, 0.5006, 29.3)	5G 6/8
9	(0.1667, 0.4869, 29.3)	5G 6/4
10	(0.1415, 0.4180, 11.7)	5B 4/6
11	(0.1577, 0.4361, 11.7)	5B 4/4
12	(0.1769, 0.4537, 11.7)	5B 4/2
13	(0.2537, 0.3307, 11.7)	5P 4/16
14	(0.2331, 0.3794, 11.7)	5P 4/10
15	(0.2132, 0.4297, 11.7)	5P 4/4
16	(0.1978, 0.4683, 100.0)	Ñ10

Experiment 1

Stimuli and Observer Task

The objective of the first experiment was to collect G_0 data for those selected chromaticities under a fixed adaptation. The stimulus background's peak luminance (nominally L^* of 100) was at three different levels of 50, 100, and 200 cd/m^2 which covered some typical color space standards. The luminances of other random neutral pixels were accordingly changed to the specified L^*. Each observer repeated the first experiment task under the three backgrounds in random order. Most of the results below will mainly focus on the 200 cd/m^2 peak background for brevity and the comparisons across different backgrounds will also be discussed.

The stimulus only had the freedom of changing luminance while its chromaticities stayed constant. The observer used a keyboard to adjust the luminance. At each trial, the stimulus started with either the lowest or the highest luminance within the gamut (for the 200 cd/m^2 peak background; for the other two backgrounds of 50 and 100 cd/m^2 the high starting point was scaled to a half and a quarter of the gamut limits respectively so that the stimuli did not appear over-bright while still glowing and the observer could lower the luminance more efficiently). The adjustments steps were set in a similar way. For the 200 cd/m^2 peak background, fine step was 5 cd/m^2 and large step was 30 cd/m^2. According to post-experiment interviews, observers did not have trouble with the steps provided.

Under each background, after a two-minute adaptation, the observer was asked to "adjust the brightness of the center patch until it just appears no grayness, or equivalently just about to glow or cease to glow depending on the starting point". The concept of grayness was illustrated by first increasing the luminance of the neutral (stimulus #16), which appeared from black to gray, and gradually to white with no grayness and to the glowing state. For other chromaticities as well as the opposite adjustment direction the demonstration was similarly repeated, and the observers were expected to learn to generalize grayness to those non-neutral colors. The observer was suggested to "adjust along either increasing or decreasing brightness direction" as much as they could. This constraint was similar to that typically used in the method of limits and adopted here because the adjustment direction might have a temporal effect on the adaptation, which we hoped to average out. Observers were also allowed to reverse back with a step of 90 cd/m^2 (for the 200 cd/m^2 peak background), especially when they passed their G_0 thresholds. Both up and down directions were repeated three times thus $16*6=96$ trials, in a random order, were done for each background. A training session covering all chromaticities and both adjustment directions were provided and the observer could practice until they felt confident about the task. There was no time limit for each trial and it took the observers on average about 1 hour for each background.

Twelve observers (7 M & 5 F; average age of 31) with normal color vision completed this experiment. 11 of them, including two of the authors, had a color science background and experience with psychophysical experiments. One naive observer participated and achieved a similar level of repeatability.

Results: G_0 and Intra- & Inter-observer Variations

Table 2 lists the G_0 luminance results averaged over both repeats and observers as well as intra-observer variations. For each observer & chromaticity combination (6 repeats), the intra-observer variations are quantified by the standard deviation (std.) and coefficient of variation (CV), which are averaged over observers in Table 2 to mainly show the chromaticity dependency. CV values, by normalizing the standard deviation to the means, are more consistent across chromaticities. The intra-observer variation is considered large relative to typical brightness matching experiments. However, as will be shown, our designs of three repetitions along both up and down direction seems to cover the reasonable variation and converge the average results well.

Table 2. Average G_0 results under 200 cd/m^2 peak background and intra-observer variations quantified by the standard deviation (std.) and coefficient of variation (CV) averaged across observers.

Stimulus	Mean G_0 (cd/m^2)	Mean std. (cd/m^2)	Mean CV
1	72.94	23.10	0.29
2	179.31	45.48	0.27
3	338.67	79.79	0.27
4	361.72	95.63	0.31
5	435.82	119.89	0.33
6	431.28	130.43	0.35
7	334.39	105.97	0.34
8	387.97	132.73	0.36
9	466.44	161.93	0.36
10	310.11	92.92	0.34
11	379.57	93.53	0.30
12	448.53	134.32	0.32
13	181.88	54.46	0.34
14	275.61	75.93	0.31
15	411.25	106.99	0.30
16	530.00	137.01	0.28

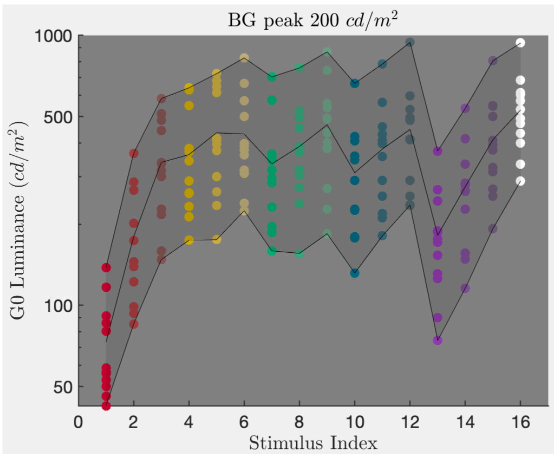

Figure 3. G_0 luminances under 200 cd/m^2 peak background. Each dot represents one individual observer's averaged result. The three connected lines correspond to the upper and lower boundaries as well as the average of all observers, which is the second column in Table 2.

In Fig. 3, each dot represents the G_0 luminance of the stimulus indexed on the x-axis for each observer. Thus, the vertical range of the shading area is the inter-observer variation. Interestingly, the variation is very similar on the log scale (the maximum to minimum ratio ranges from 3.2 to 5.0). This plus the CV values previously likely reflects the deficit of (linear and absolute) luminances across different chromaticities. The three connected lines in Fig. 3 correspond to the upper and lower boundaries as well as the average of all observers, which is the second column in Table 2. Although none of them exactly correspond to an individual observer, the shape or general trend is shared within individual observers. Such trend is consistent with the H-K effect which is a function of both hue and chroma. Red and purple have a stronger H-K effect so that they require lower luminances to be equally bright. For a given hue, higher saturation appears brighter when iso-luminant to the neutral thus need less luminance at constant brightness level. The difference between equal luminance and equal brightness is accounted for in Evans's chromatic strength concept, which aligns with our results.

The aggregated results along either up or down adjustment direction are plotted in Fig. 4. The average of the two lines would be the center line in Fig. 3. It can be seen that the "down" direction led to a higher G_0 threshold. Since it started with a high luminance start point, it was likely that the adaptation depended more on a higher average luminance of what the observer had seen. The difference between the two lines again can be explained by chromatic strength: the saturated red and purple had a higher multiplying factor in luminance-to-brightness conversion so that the observers were more sensitive.

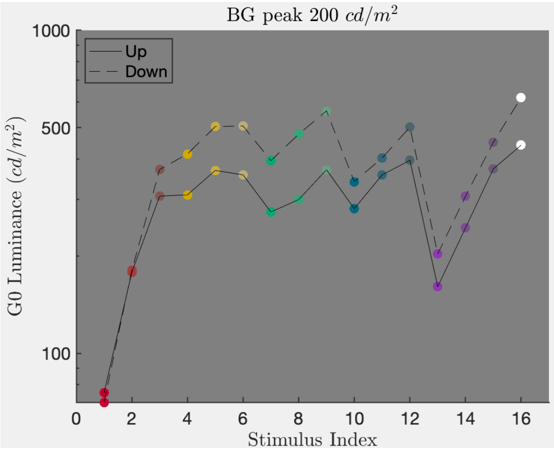

Figure 4. *G_0 luminances under 200 cd/m^2 peak background via different adjustment directions.*

Results: G_0 under Different Adaptation Backgrounds

Figure 5 shows the G_0 under different adaptation backgrounds, plotted as peak 100 or 200 cd/m^2 versus peak 50 cd/m^2. They both exhibit high linear correlations ($r = 0.9835, p < 0.001$ and $r = 0.9842, p < 0.001$, respectively). However, the linearity did not follow the Y/Y_n invariant, expected to follow the two dash lines, as G_0 under each adaptation was supposed to have same lightness [17]. The relation between the background used in the experiment (or more generally any background) and a computational Y_n need further investigations.

Experiment 2

Stimuli and Observer Task

The stimulus configuration was similar to that shown in Fig. 1. Two 3-by-3 deg color patches were similarly placed in the display center with a separation of one deg between them. The collected mean G_0 data under the 200 cd/m^2 peak background from the first experiment (Table 2) was used to test whether they would appear equally bright when seen side by side.

A paired comparison (2-alternative-forced-choice) between the 16 G_0 colors was done by 12 observers with normal color vision (7 M & 5 F; average age of 32), among which eight observers did the first experiment and three were naive observers with no color science background. The task was to select "which patch (left or right) appears brighter". Before the experiment, a brief training session was given where the formal definition of brightness and some examples of colors with different hue/saturation and luminance (G_0 or half G_0) were pre-

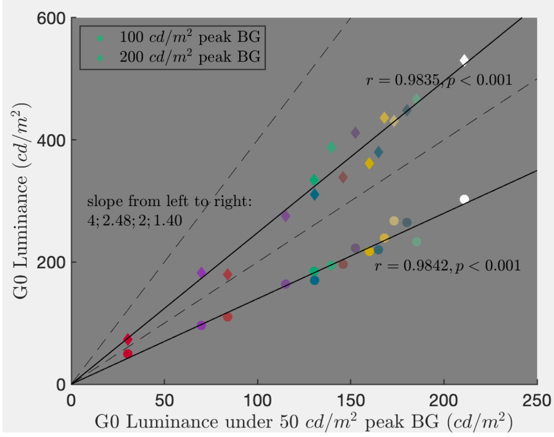

Figure 5. *G_0 luminances under different backgrounds. The dash lines serve as reference if there is linear scaling between different backgrounds. The Pearson correlation coefficients (r) are indicated for those G_0 across different adaptations.*

sented. The observers went through a 2-min adaptation to the background. All G_0 compared to each other in terms of brightness resulted in $16*15/2 = 120$ trials, which were repeated four times. In each block, the 120 trials were randomized. Since G_0 colors were supposed to be equally bright, which would make the experiment difficult, extra trials where two patches had the same chromaticity with either G_0 and half G_0 luminance were added. All 16 chromaticities were repeated three times to gauge the observer consistency. In all trials, the left and right order were also randomized. Three observers only made one wrong choice (of selecting half G_0 as brighter) in the 48 extra trials, and the rest of the observers made all correct decisions. In addition, another set of G_0 and half G_0 comparisons between different chromaticities was supplemented (only 10 low- & middle-saturation colors plus neutral, pairwise compared in 55 trials). Those G_0 vs. half G_0 checking comparisons were not used in deriving the Thurstonian scales below.

Results: Brightness Scale Derived from Paired Comparisons

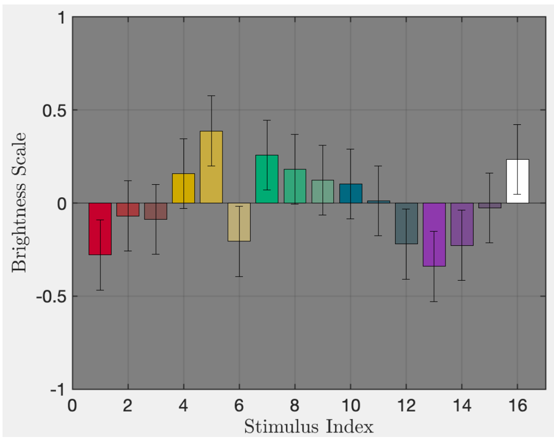

Figure 6. *Thurstonian brightness scale derived from all observers' paired comparisons.*

With the Thurstone Case V assumption, the comparison results were converted to an interval scale shown in Fig. 6 [21]. The 95% confidence interval was calculated using an empirical formula [22], where the number of observations N was set to the number of observers, i.e., 12. On the y-axis, zero indicates the mean value and one corresponds to one unit of standard deviation. The ideal case, if all stimuli at G_0 appear equally bright, would be all scale values converge to zero. Although not perfectly, the results show that most of them are not significantly different from each other, with the maximum deviation from the mean is still within ~0.4 standard deviation. It can be found that most pairs of colors have very similar scale values, especially those sharing similar hue and saturation. The two saturated red and purple had the least scale value, which means the G_0 collected from the first experiment was lower than what this group of observers needed. This might also come from the trade-off between being bright and saturated and the difficulty of hetero-chromatic brightness matching (particularly when there is no fixed reference). One (naive) observer commented that she was not sure whether brighter meant more whitish or more glowing.

Among the 55 trials of G_0 and half G_0 comparisons between different chromaticities, four observers (three of them were not included and one was the only naive observer in the first experiment) made more than three selections towards the half G_0 patch which had lower saturation. It is possible that they did not consider all the chromatic strength as part of brightness. When excluding the four observers, more scales became closer to zero, as shown in Fig. 7. In particular, when comparing to the traditional brightness matching reference, stimulus #16, only stimulus #13 is significantly different. Note the paired comparison experiment forced the observers to make discriminatory decisions, so that the statistical significance does not necessarily translate to significant perceptual difference.

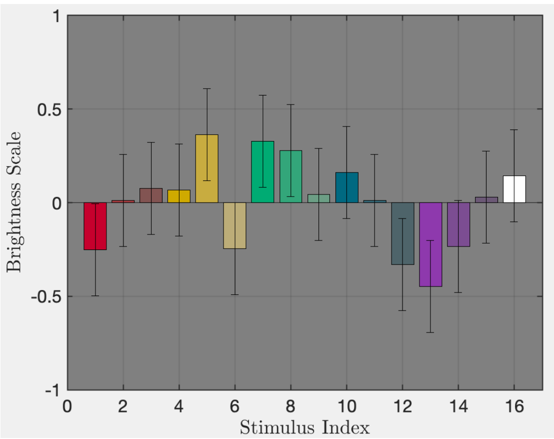

Figure 7. Thurstonian brightness scale derived from the paired comparisons of all observers except those observers who might confuse brightness with lower saturation.

Discussion

From the two experiments, the collected G_0 seems a promising result of equally bright colors that incorporate the H-K effect. The relation between G_0 lightness/brightness threshold and appearance mode boundaries has been discussed by Evans [9] and others [23, 14], where the MacAdam's optimal colors were considered to determine the physical constraint of the observer's prior of whether a color appears reflective or self-luminous. Fairchild and Heckaman suggested using NCS zero blackness as a computational proxy for G_0 [17]. Figure 8 presents the relations between the collected G_0 luminance and the two physical gamut boundaries, which have high correlations ($r = 0.9390, p < 0.001$ and $r = 0.9076, p < 0.001$, for optimal colors and NCS zero blackness, respectively). The optimal colors usually have slightly higher luminances than NCS zero blackness. And there is a ratio of ~2.6 between the visual results and the physical results. Perfectly diffuse white has a luminance of 200 cd/m^2 under a 200 cd/m^2 illumination whereas the neutral had a higher luminance to be G_0 under our 200 cd/m^2 peak background. The equivalence between the background and illumination is a similar problem that has been discussed in the across-adaptation results.

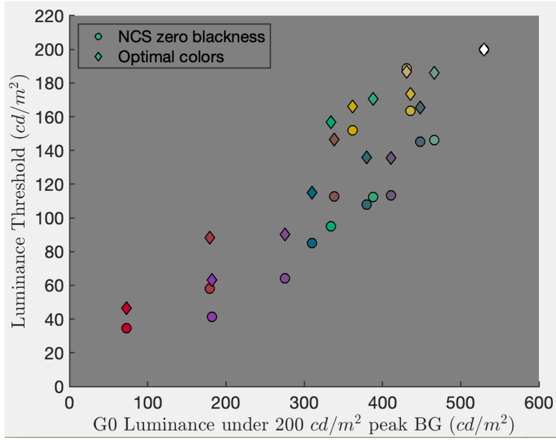

Figure 8. G_0 under 200 cd/m^2 peak background versus NCS zero blackness and the optimal colors' luminance under a 200 cd/m^2 illumination.

The physical or computational boundaries have no uncertainties, however, the visual boundaries exhibited high intra- & inter-observer variations. While it seems at least the repeats from both up and down directions helped constrain the average results to be reasonable, those variations indicated both the difficulty of judging G_0 or luminosity and the individual difference. Previous studies reported different ratios between G_0 or luminosity threshold and the background/illumination [24, 23] and high inter-observer variations in a similar task [25]. In particular, stimulus #16 at theoretical G_0 approximately corresponds to the practical reference point (Y_n) in CIELAB and CIECAM02. As mentioned in the Introduction section, Y_n is more well-defined for illuminated physical objects with a corresponding physical identity (e.g., the perfectly reflective diffuser) than the display rendering counterpart, which has more flexibility in manipulating luminance (relations) across the screen. More experiments are needed to find the determinants of the variations and devise better psychophysical approaches. Although a brightness matching experiment may help reduce the inter-observer variations, the appearance equivalence will only be achieved across chromaticities but not across observers. For a given observer, they might agree that the two colors appear equally bright but not necessarily at zero grayness or glowing state in an absolute sense. Our results based on each observer's internal absolute reference also achieve equal brightness across chromaticities, which can be served as reference anchors to derive uniform lightness/brightness scales.

The intra- & inter-observer variations also have implications on how to use descriptive statistics from the results. In partial hindsight, as luminance on the log scale seems better than linear luminance, the average G_0 across either trials or observers might be better calculated using geometric means instead of arithmetic means. Both as well as the median were compared; the geometric means are lower than the arithmetic means as expected, mostly around ~15%, and the medians are relatively closer to the arithmetic means. Since our second experiment was based on the arithmetic means and achieved promising results for most observers, we would only present and stay with the current results and share our complete raw results in the future. Also, for a similar reason that we did not assume any potential models before the experiments, the adjustment steps in the first experiment were based on linear luminance units. For future experiments, we might consider the log unit as well.

Conclusion

In this work, two psychophysical experiments were conducted based on the concepts of zero grayness (G_0) and brilliance. In the first experiment, the G_0 results for 16 representative chromaticities were collected as a threshold with equal brightness under three different adaptation backgrounds. Our visual results, consistent across observers, align with the luminance versus brightness discrepancy and the H-K effect and show good correlations across adaptation levels and against two physical gamut boundaries, i.e., the MacAdam's optimal colors and NCS zero blackness. The hypothesis that those G_0 colors should appear equally bright was verified by a second experiment, where another group of observers did paired comparisons between those G_0 colors. Our ongoing and future work involves using those anchors to derive uniform lightness/brightness scales.

Acknowledgments

We are grateful to all the observers for their inputs and to the anonymous reviewers for their comments.

References

[1] M. D. Fairchild, *Color Appearance Models*, 3rd ed., Wiley-IS&T Series in Imaging Science and Technology. Chichester, West Sussex: Wiley, 2013.

[2] P. Lennie, J. Pokorny, and V. C. Smith, "Luminance," *Journal of the Optical Society of America A*, vol. 10, no. 6, p. 1283, 1993.

[3] J. Koenderink, A. van Doorn, and K. Gegenfurtner, "Color weight photometry," *Vision Research*, vol. 151, pp. 88–98, 2018.

[4] P. K. Kaiser and G. Wyszecki, "Additivity failures in heterochromatic brightness matching," *Color Research & Application*, vol. 3, no. 4, pp. 177–182, 24.

[5] M. Ayama and M. Ikeda, "Brightness-to-luminance ratio of colored light in the entire chromaticity diagram," *Color Research & Application*, vol. 23, no. 5, pp. 274–287, 1998.

[6] M. D. Fairchild and E. Pirrotta, "Predicting the lightness of chromatic object colors using CIELAB," *Color Research & Application*, vol. 16, no. 6, pp. 385–393, 1991.

[7] S. Hermans, K. A. G. Smet, and P. Hanselaer, "Color appearance model for self-luminous stimuli," *Journal of the Optical Society of America A*, vol. 35, no. 12, p. 2000, 2018.

[8] R. M. Evans and B. K. Swenholt, "Chromatic Strength of Colors: Dominant Wavelength and Purity," *Journal of the Optical Society of America*, vol. 57, no. 11, p. 1319, 1967.

[9] R. M. Evans, "Fluorescence and Gray Content of Surface Colors," *Journal of the Optical Society of America*, vol. 49, no. 11, p. 1049, 1959.

[10] Y. Nayatani, H. Sobagaki, and K. Hashimoto, "Relation between Helmholtz-Kohlrausch Effect, Purity Discrimination, and G_0 Function." *Journal of Light & Visual Environment*, vol. 17, no. 2, pp. 16–24, 1993.

[11] Y. Nayatani and H. Sakai, "A relationship between zero-grayness luminance by Evans and perceived brightness of spectrum colors," *Color Research & Application*, vol. 33, no. 1, pp. 19–26, 2008.

[12] R. L. Heckaman, "Brilliance, contrast, colorfulness, and the perceived volume of device color gamut," Ph.D. dissertation, Rochester Institute of Technology, 2008.

[13] R. L. Heckaman and M. D. Fairchild, "G_0 and the gamut of real objects," in *Proceedings of AIC Color*, 2009.

[14] J. M. Speigle and D. H. Brainard, "Luminosity thresholds: Effects of test chromaticity and ambient illumination," *Journal of the Optical Society of America A*, vol. 13, no. 3, p. 436, 1996.

[15] T. Fujine, T. Kanda, Y. Yoshida, M. Sugino, M. Teragawa, Y. Yamamoto, and N. Ohta, "Relationship between mode-boundary from surface color to fluorescent appearance and preferred gamut on wide-gamut displays," *Journal of the Society for Information Display*, vol. 18, no. 8, p. 535, 2010.

[16] M. Bertalmío, *Vision Models for High Dynamic Range and Wide Colour Gamut Imaging: Techniques and Applications*, Computer Vision and Pattern Recognition. Academic Press, 2019.

[17] M. D. Fairchild and R. L. Heckaman, "Deriving appearance scales," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2012, pp. 281–286.

[18] M. D. Fairchild, "A digital test chart for visual assessment of color appearance scales," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2021, *in press*.

[19] E. A. Day, L. Taplin, and R. S. Berns, "Colorimetric characterization of a computer-controlled liquid crystal display," *Color Research & Application*, vol. 29, no. 5, pp. 365–373, 2004.

[20] M. D. Fairchild, "A victory for equivalent background—on average," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 1999, pp. 87–92.

[21] P. G. Engeldrum, *Psychometric scaling: a toolkit for imaging systems development*. Imcotek Press, 2000.

[22] E. D. Montag, "Empirical formula for creating error bars for the method of paired comparison," *Journal of Electronic Imaging*, vol. 15, no. 1, p. 010502, 2006.

[23] K. Uchikawa, K. Koida, T. Meguro, Y. Yamauchi, and I. Kuriki, "Brightness, not luminance, determines transition from the surface-color to the aperture-color mode for colored lights," *JOSA A*, vol. 18, no. 4, pp. 737–746, 2001.

[24] F. Bonato and A. L. Gilchrist, "Perceived area and the luminosity threshold," *Perception & Psychophysics*, vol. 61, no. 5, pp. 786–797, 1999.

[25] C.-W. Lin, P. Hanselaer, and K. A. Smet, "Relationship between perceived room brightness and light source appearance mode in different media: reality, virtual reality and 2d images," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2020, pp. 30–35.

Author Biography

__Hao Xie__ is a PhD student in the Munsell Color Science Lab at Rochester Institute of Technology. He received his B.Eng. degree in optical engineering from Zhejiang University, China, and his M.S. degree in electrical engineering from the University of Rochester. His research interests include applied vision science and computational photography.

__Mark D. Fairchild__ is Professor and Founding Head of the Integrated Sciences Academy in RIT's College of Science and Director of the Program of Color Science and Munsell Color Science Laboratory. He received his B.S. and M.S. degrees in Imaging Science from R.I.T. and Ph.D. in Vision Science from the University of Rochester. Mark was presented with the 1995 Bartleson Award by the Colour Group (Great Britain) and the 2002 Macbeth and 2018 Nickerson Awards by the Inter-Society Color Council (ISCC) for his works in color appearance and color science and contributions to the field. He is a Fellow of the Society for Imaging Science and Technology (IS&T) and the Optical Society of America (OSA). Mark was also presented with the Davies Medal by the Royal Photographic Society for contributions to photography, the 2008 IS&T Raymond C. Bowman award for excellence in education, and the 2021 SID Otto Schade Prize for contributions to perceptual aspects of image quality.