Received: 18 August 2022 | Revised: 25 November 2022 | Accepted: 13 December 2022
DOI: 10.1002/col.22839

RESEARCH ARTICLE

COLOR RESEARCH AND APPLICATION WILEY

The Helmholtz-Kohlrausch effect on display-based light colors and simulated substrate colors

Gregory High | **Phil Green** | **Peter Nussbaum**

Norwegian Colour and Visual Computing Laboratory, Norwegian University of Science and Technology, Gjøvik, Norway

Correspondence
Gregory High, NTNU in Gjøvik, Postboks 191, NO-2802 Gjøvik, Norway.
Email: gregory.high@ntnu.no

Funding information
Norwegian University of Science and Technology

Abstract

The Helmholtz-Kohlrausch (H-K) effect is investigated in relation to light colors of every hue, including those typical of print substrate colors that might be simulated on a graphic arts display. A method of adjustment is used in conjunction with a soft-proof setup, in which an achromatic stimulus is adjusted until it has the same lightness appearance as a set of test colors. Higher chroma colors are found to appear lighter than their metric L^* would indicate. The H-K effect is found to be quite strong in bluish colors, but negligible in yellowish colors, consistent with several previous studies. However, qualitative analysis reveals a peak H-K effect in red-magenta hues. We propose a modification to Fairchild and Pirrotta's existing H-K lightness appearance function[1] which addresses the peak in red hues, and which may prove beneficial in hue-dependent applications.

KEYWORDS

brightness matching, Helmholtz-Kohlrausch, hue dependency, lightness matching, soft proofing

1 | INTRODUCTION

First attributed to Helmholtz,[2] the contribution of colorfulness to brightness appearance has long been recognized. It is evident in side-by-side heterochromatic comparisons, where colors appear brighter as they become more saturated. This is particularly noticeable when comparing high chroma colors to a reference achromatic color of equal luminance. This phenomenon is referred to as the Helmholtz-Kohlrausch effect.[3,p.121]

Generally, more saturated colors require less luminance (or a lower reflectance factor) than neutral colors in order to appear equally bright (or light). The H-K effect may thus be thought of as a failure of additivity. A visual agreement between equal-luminance and equal-perceived lightness is achievable, but only under specific viewing conditions, such as the flicker-photometry used to generate the CIE luminous efficiency $V(\lambda)$ function.[4,p.263]

Sanders and Wyszecki[56] determined a lightness index (L) for color samples, which is the reflectance factor of an achromatic reference sample that has the same lightness appearance as the test color. Thus, samples with equal lightness appearance (and equal L), but differing in reflectance factor (Y), could be described using L/Y ratios.

In an experiment in the object viewing mode,[5] 37 Munsell papers, all of similar luminous reflectance, were evaluated in terms of lightness against reference colors in the Munsell neutral scale. The L/Y ratios were fitted to the XYZ tristimulus values of the samples using a second order polynomial with cross-products. Despite the limited gamut of the training set, the fitted model described the H-K effect as being weakest in the yellow hues, and increasing in strength towards blue hues. However, the model also showed a steep increase in H-K towards the red-magenta hues, and the authors extrapolated that more saturated magentas would require a far lower luminance than yellow hues in order to have the same lightness appearance.

In a further experiment, Sanders and Wyszecki[6] investigated with many more samples of varying reflectance. Notably, the L/Y ratio at each CIE chromaticity co-ordinate held close to constant across a wide range of different reflectance factors (up to an L^* of approximately 80).

This concept may be carried forward to work in a uniform color space. In order to predict the H-K effect, a general approach is to add some fraction of colorfulness to the achromatic intensity in order to achieve a better prediction of brightness for more saturated colors. Chroma is defined as a measure of "colourfulness of an area judged as a proportion of the brightness of a similarly illuminated area that appears grey [or] white".[7] In this way, chroma will remain constant for related colors when viewed or measured under different absolute levels of illumination. Therefore, we see that adding a factor of a sample's chroma to its metric lightness may be expected to predict the equivalent achromatic lightness, consistent with the H-K effect in heterochromatic surface colors. This takes the form of Equation (1), where L^* and C^*_{ab} are the colorimetric values of a sample, k is a constant (some factor of C^*_{ab}), and L^*_{EAL} is the predicted equivalent achromatic lightness.

$$L^*_{\mathrm{EAL}} = L^* + k \cdot C^*_{ab} \quad (1)$$

However, this approach does not predict the hue angle dependency of the H-K effect, particularly for yellow hues where the effect is very limited.

To address this, Fairchild and Pirrotta[1] produced a model that implemented a ½-cycle sine-based function relative to a hue angle of 90° (yellow), which in turn predicted a maximal effect in the complementary blue hue. The model returned a modified L^* metric from the original CIELAB colorimetry. In addition, a new set of observer data was produced using a method of adjustment, whereby a simulated achromatic surface reflectance was adjusted to match the perceived lightness of a series of physical color samples. The inclusion of a multiplicative factor for hue-angle dependency was found to improve the fit of their model systematically.

Subsequent CAMs for surface colors have rarely included a predictor for the H-K effect, since the calculations add complexity and are not easily invertible, and are very dependent on the angular subtense of the stimulus. The more comprehensive CAM97c does model the H-K effect, and includes predictors of lightness and brightness modified with the ½-cycle sine-based function described above.[8]

Recent CAMs for unrelated colors have revisited the H-K effect. Several of these adopt the general approach described above of computing brightness as the sum of the achromatic (luminance-derived) brightness and some optimized constant of colorfulness.[9] This does improve the overall fit to the perceptual data. However, the hue-dependent function developed by Fairchild and Perrotta has seen further application, most notably in CAM15uz.[10] For this purpose, Huang et al.[10] created a new set of observer data for unrelated colors, using color stimuli across a wide range of chromaticities specifically to investigate the H-K effect. With coefficients that were optimized to fit their data set, the revised model with the hue-dependent function found much improvement over their previous CAM15u. However, hue dependent H-K is not the only consideration, and Cam20U[9] improves performance further without a hue-dependent function, instead including parameters for rod receptor intrusion (known to affect colors in the mesopic range), and the angular subtense of the stimulus.

Recent work by Xie and Fairchild[11] used Evans's notion of G_0 (zero grayness). Under a fixed adaptation, the luminance levels of a series of chromatic stimuli were adjusted to derive a boundary between surface mode colors and aperture mode colors (below which colors appeared to have some grayness, and above which colors appeared to glow), though both were simulated on a display. The hue-dependent nature of this boundary was consistent with the H-K effect, and also featured a peak in red hues, and to a lesser extent in purples.

There is currently renewed interest in modelling the H-K effect. Park[12] has investigated its impact on the very saturated colors that can be generated using wide-gamut and high-luminance OLED displays. Hellwig et al.[13] have revisited several previous datasets, and propose a new extension to their refinement of CAM16.[14]

In the previous phase of our work[15] we investigated the H-K effect in near-white substrate colors, simulated on a display and consistent with graphic arts viewing conditions. Only six hue angles were selected for investigation, with 90° (yellow) and 270° (blue) expected to provide the minima and maxima respectively for the H-K effect. However, a strong H-K effect in red and magenta hues was observed. This demonstrated that far greater granularity was required to describe anything approaching a continuous hue dependent function.

The present phase of work has therefore been undertaken to ascertain the hue-dependent nature of the H-K

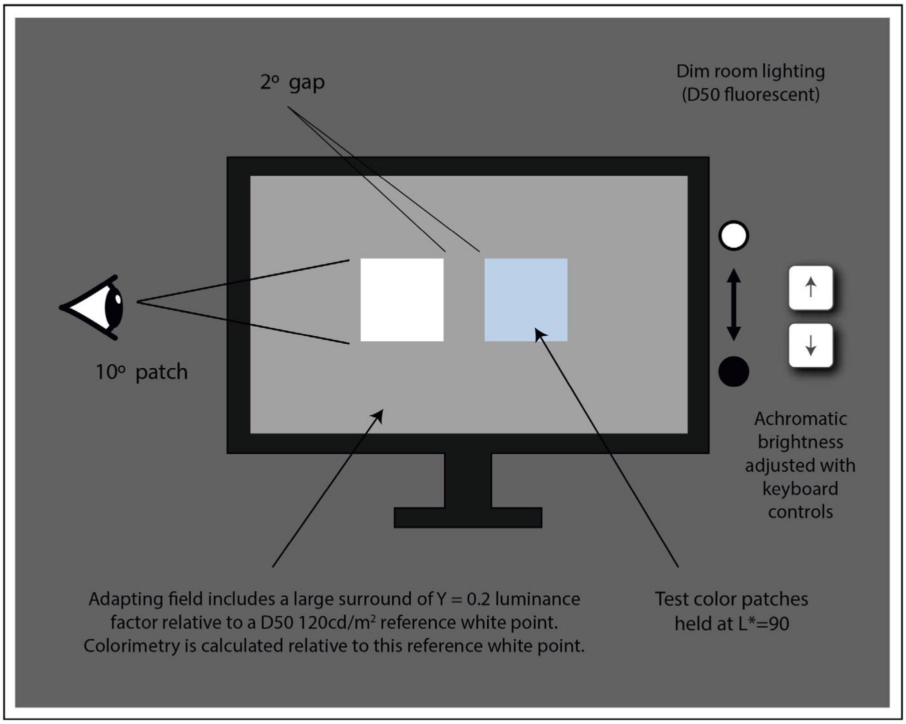

FIGURE 1 User interface presented on a 24″ Eizo CG248-4K, consistent with P2 viewing conditions

effect amongst light and near-white colors in a typical graphic arts soft-proofing setup. The aim of the present work is to take a systematic approach to lighter colors of every hue, to determine the distribution of brightnesses that describe the H-K effect, and to model the effect for these particular viewing conditions.

2 | METHOD

Following the method described in a previous paper,[15] an Eizo CG248-4 K graphic arts display was calibrated and profiled with Eizo's ColorNavigator software, and using a Konica Minolta CS2000 TSR. The TSR gave the benefit of including any reflected ambient light in the measurements from the position of a seated observer. The display was prepared to a D50 white point at a target luminance of 200 cd/m^2.

A color-managed interface was developed using Matlab, and a full-screen gray background was set at 24 cd/m^2. Target colorimetry was then scaled to an adopted white point of D50 at 120 cd/m^2, in accordance with P2 viewing conditions described in ISO 3664.[16] In this way the hardware's native white point was hidden from view, and the extra “headroom” in luminance allowed for color stimuli which would normally be clipped by a traditional display setup. The combined hardware and user interface was checked for accuracy, with 50 test patches measured independently with the TSR and compared to their target colorimetry. All test colors were within the gamut of the display, with a mean error of 0.43 ΔE_{00}, and a maximum of 1.16 ΔE_{00}. Both the accuracy and uniformity of the combined system was found to be well within the recognized tolerances for soft-proofing described in ISO 14861.[17]

2.1 | User interface

A full-screen user interface was presented, featuring a uniform gray surround with a luminance factor of 0.2 relative to the specified 120 cd/m^2 white point luminance. Reference and test patches were then positioned side-by-side with a 2° gap, each patch subtending a viewing angle of 10° from the viewing position (see Figure 1 for details).

As before,[15] this side-by-side comparison on a mid-gray background is typical of print-to-proof appearance comparisons for different simulated substrates. Also in line with a graphic arts approach, we use D50 colorimetry utilizing the CIE 1931 2° Standard Observer throughout.

2.2 | Method of variable-achromatic-color (VAC)

Two methods of adjustment are recommended by Nayatani[18,p.386]: a method of variable-achromatic-color (VAC) (adjusting an achromatic stimulus to match the brightness of a chromatic test color); and a method of variable-chromatic-color (VCC) (adjusting a chromatic stimulus to match the brightness of an achromatic test color). In our previous work we found that the H-K effect was reported as stronger in low chroma colors when using VCC. However, VAC was considered to be an easier task for observers, and returned results with lower variability. For this new phase of work we therefore adopted the VAC method of adjustment.

For each chromatic test color an achromatic patch was adjusted via the keyboard controls to make a lightness appearance match. The starting point of the achromatic patch was always L^*-20 darker than the test color, to avoid adverse adaptation problems or potential display gamut clipping issues. The observer was then free to increase or decrease the CIELAB L^* (in the lightness dimension only), iterating until a lightness appearance match to the test patch was obtained.

The presentation order and left/right patch positions were randomized. In addition, observers were free to adjust the magnitude of the adjustment in order to fine-tune their lightness match.

A short training session ensured observer adaptation, and directions were given to observers both verbally and in written form. Observer details were captured using an input dialogue box, and in accordance with local GDPR requirements these were anonymized at the point of collection.

2.3 | Test colors at all hue angles

To be representative of light colors at all hue angles, all test colors were selected with a metric L^* of 90. Test colors were selected along 24 hue angles (at an h_{ab} interval of 15°). For each hue angle colors were sampled at two chroma levels (with C^*_{ab} of 15 and 30). These combinations include a small number of colors repeated from the previous work.[15] Based on this previous experience, the H-K effect would not be expected to increase the apparent lightness of these test colors (metric L^* of 90) beyond that of an achromatic patch (with an L^* of 100). However, as is the case in soft-proofing of prints, the simulated substrate color will be the brightest color in the field of view.

2.4 | Observers

Twenty-one observers took part (14 male, 7 female) aged from 21 to 59 with an average of 29.7 years. All were known to have color-normal vision.

3 | RESULTS

Test colors were of various hue angles and chroma levels, but always a metric L^* of 90. For each test color we obtained the average matching observer-adjusted achromatic lightness L^*_O. Results are plotted in Figure 2.

The hue dependency of the H-K effect is seen to be very similar for colors with a C^*_{ab} of 15 and 30, differing mainly in magnitude. The present results are broadly in line with the VAC results obtained previously for the six hue angles (h_{ab} of 30°, 90°, 150°, 210°, 270°, and 330°),[15] with the exception of the green hues at 150° where the H-K effect was previously judged to have been less pronounced. A steep trough is observed in the yellow region (around an h_{ab} of 90°), where the effect is minimal for the samples with C^*_{ab} of 30, and practically null for those with C^*_{ab} of 15. The distribution is approximately bimodal, with a rough plateau in the green-cyan-blue regions (from around an h_{ab} of 150° through to an h_{ab} of 270°). A distinct peak is observed in the magenta-red regions (from around an h_{ab} of 300° through to an h_{ab} of 30°).

Inter-observer variability was calculated with a mean standard deviation of 5.17 for C^*_{ab} of 15, and 5.99 for C^*_{ab} of 30.

The most noticeable outlier is for the C^*_{ab} of 15 patch at h_{ab} of 0°\360° (a magenta-red hue), where the mean observed lightness is far lower than for adjacent hue angles. Otherwise, the collected data points give the impression of a near-continuous function across all hues. A trend line is therefore added to the data, derived from a 3 point moving average, which provides a smoother visual guide to the hue dependent behavior.

The same data and trend lines are re-plotted as polar co-ordinates in Figure 3. The radial values represent the observed lightnesses of colors ranging from L^*_O of 90 to 105, which are greater than the stimuli's colorimetric L^* of 90, depending on hue angle. The resultant shapes are similar to the ½-cycle sine-based distribution described elsewhere.[1] However, the plot reveals the peak H-K effect in the red-magenta region, and this gives the distribution an irregular "apple-shaped" appearance.

4 | MODELLING THE H-K EFFECT

Visual inspection of the data plotted in Figure 3 reveals a distribution with two axes along which the H-K effect may be seen to increase: the yellow-to-blue (b^*) axis and also the green-to-red (a^*) axis. The change in magnitude appears greatest around the yellow hues (the trough around an h_{ab} of 90°, ±60° approx.). A change in

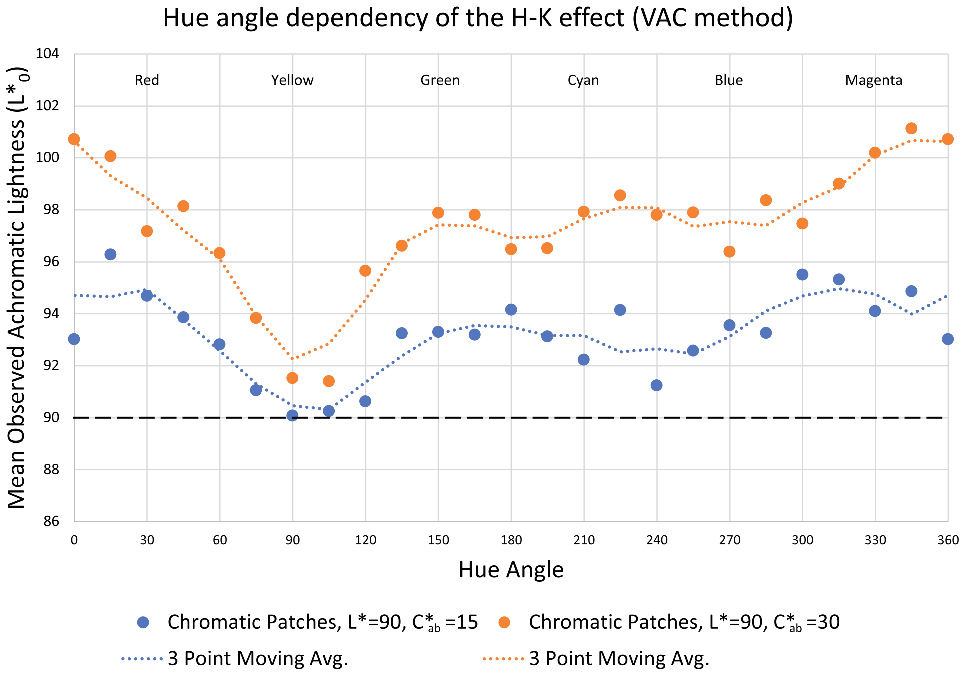

FIGURE 2 Mean observed achromatic lightness (L^*_O) of test colors across all hue angles. Mean inter-observer standard deviation is 5.58. All test colors are held at L^* of 90

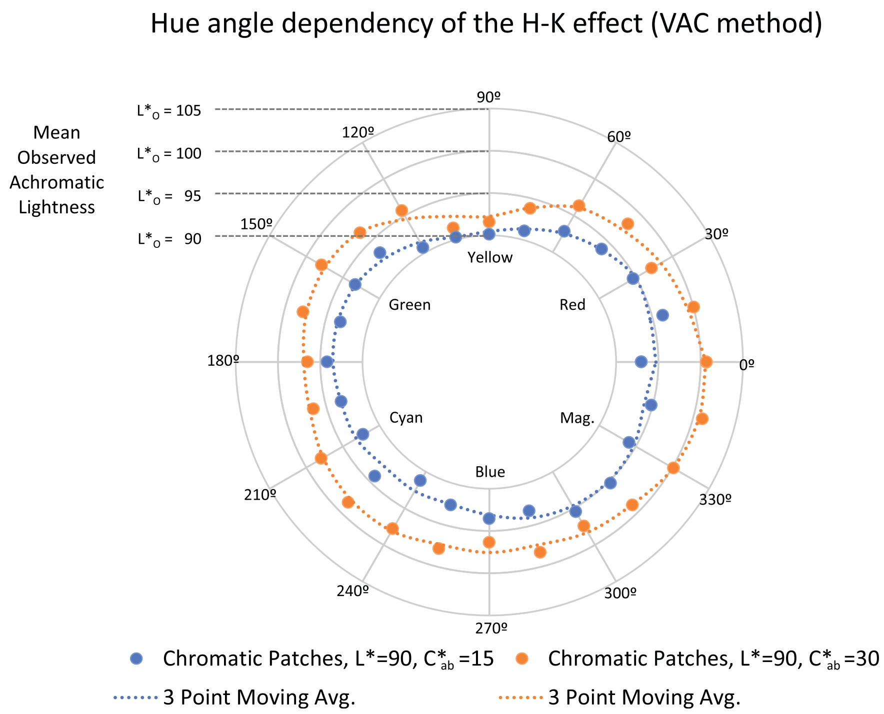

FIGURE 3 Mean observed achromatic lightness (L^*_O) of test colors of different hue, represented using polar co-ordinates. All test colors are held at L^* of 90

magnitude is also noticeable around the peak in red hues (an h_{ab} of 0°, ±60° approx.)

A model should be able to predict a scaling factor for chroma at each hue angle, which returns an equivalent achromatic lightness (L^*_{EAL}) consistent with the H-K effect. By rearranging Equation (1) we can see that an indicative $k_{(h^\circ)}$ can be derived directly from observer data (see Equation (2), where L^*_O is the observed lightness, L^* is colorimetric lightness, C^*_{ab} is chroma, and $k_{(h^\circ)}$ is the hue-dependent scalar).

$$k_{(h^\circ)} = \frac{L^*_O - L^*}{C^*_{ab}} \quad (2)$$

4.1 | Previously published datasets

We use the approach outlined above to analyze three sets of observer data, as well as one previous model. $k_{(h^\circ)}$ is calculated for each test color from the available colorimetry, and these are arranged hue-wise. Results are plotted in Figure 4.

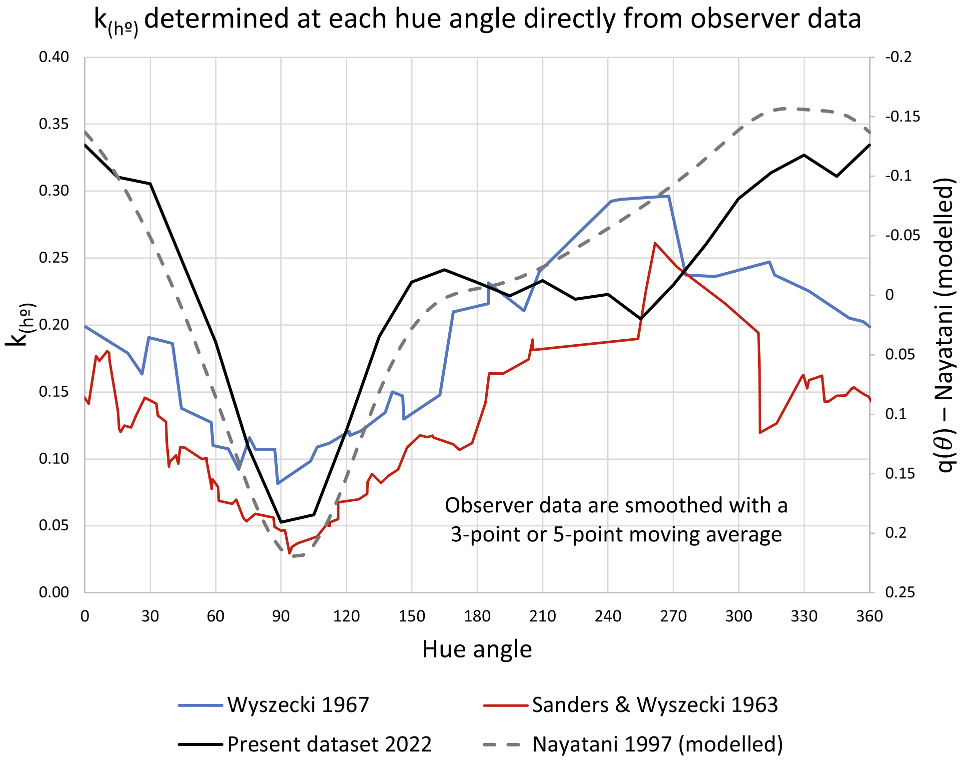

FIGURE 4 Primary ordinate axis: an indicative $k_{(h°)}$ hue-dependent scalar derived from smoothed observer data. Secondary ordinate axis: Nayatani's hue-dependent function $q(\theta)$. The two scales are not normalized

Firstly, we reference Wyszecki's 1967 surface color dataset, which was obtained by comparing chromatic ceramic tiles against achromatic reference tiles.[19] It consists of 43 samples of similar lightness (Munsell value 6) but with a good variety of hues and chroma levels. Samples are viewed on a tabletop with a matt gray background, and under daylight or simulated daylight conditions. The experiment is therefore strictly in the object mode of appearance. $k_{(h°)}$ is calculated for each sample, and the data are smoothed using a 3-point moving average. Due to the non-uniform distribution of hues in the dataset the presentation remains somewhat jagged. We see that the underlying function closely resembles a ½-cycle sine-based distribution, with the characteristic trough around h_{ab} of 90°, and a peak in the blue region around h_{ab} of 270°.

Secondly, we reference Sanders and Wyszecki's 1963 self-luminous color dataset.[20] This experiment is interesting because it uses adjustable backlit stimuli in a 10° bipartite field configuration, but with a 40° surrounding field that is illuminated from the front. The experiment consists of 96 test colors. An achromatic stimulus is adjusted until a brightness match is made with each test color, with luminance held at 20 cd/m^2 for the test color and the background. The authors debate the correct terminology to use, since some colors may appear in the object mode relative to the background, whilst colors brighter than the background may appear in aperture mode. Brightness, rather than lightness, is the preferred term, and the luminance of the matching achromatic stimulus is used as a brightness index (B).

Additionally, there is a mismatch in chromaticity between the achromatic stimulus (tungsten) and the illumination of the surrounding field (fluorescent), which causes a hue-shift in the resultant colorimetry depending on which is selected as the adopted chromaticity. Calculating the colorimetry relative to the surrounding field's chromaticity causes the trough in the graph to align with a hue angle of 90°, consistent with other datasets. We can therefore expect that the tungsten stimulus would have appeared slightly orangish relative to the background, and this may explain why the magnitude of the B/Y ratios are smaller than the L/Y ratios in the previous object mode experiment. After calculating $k_{(h°)}$ for each sample, the data are smoothed using a 5-point moving average. The non-uniform distribution of hues in the dataset, together with the sparsity of data points between h_{ab} of approximately 240° and 310°, gives the graph a rather jagged appearance. (Two samples which give extreme outliers of $k_{(h°)}$ are removed from the set: 482/1 and 475/2. These two low chroma colors demonstrate a weakness in the form of Equation 2, whereby a small over-estimate of lightness is exacerbated by a very low chroma value, giving an unusually high value for $k_{(h°)}$. Interestingly, these two sample colors are also closest in chromaticity to the surrounding field, and this may have caused some additional confusion.) Again, the underlying function resembles a ½-cycle sine-based distribution.

Thirdly, we reference the present dataset (previously plotted in Figures 2 and 3), which consists of 48 stimuli of a single lightness ($L^* = 90$) and two chroma levels (C^*_{ab} of 15 and 30) on a graphic arts display. $k_{(h^\circ)}$ is calculated for each sample, smoothed using a 3-point moving average, and the results are averaged across the two chroma levels. As we have already seen in Figure 2, whilst the present dataset shares the trough in the yellow region, the H-K effect is far stronger in the red-magenta region, where the difference between this and the previous two datasets is greatest. There is also a smaller secondary peak in the green region around h_{ab} of 150°. Whereas the red-magenta peak is consistent with other works[15,11] the smaller peak in the green hues was not so pronounced in our previous experiment[15] using exactly the same setup and method and with a proliferation of stimuli at that hue angle. Colorimetrically, all of these lighter colors lie towards the upper bounds of the gamut of real surface colors, and the strong magenta peak in the H-K effect is consistent with that found in Xie and Fairchild's G_0 *equally bright reference boundary*,[11] which may be thought to incorporate the H-K effect at the interface between (simulated) surface and aperture colors.

Lastly, we include Nayatani's 1997 hue-dependent function $q(\theta)$.[18] This function was incorporated into four separate models, predicting the H-K effect for each of the VAC and VCC brightness-matching methods, and for both object and luminous colors. A set of CIELCh coordinates (of equal lightness and chroma across all hue angles) is converted to $u'v'$ chromaticity co-ordinates relative to an Illuminant C white point, before calculating $q(\theta)$ at each hue angle. The $q(\theta)$ scale is not normalized to $k_{(h^\circ)}$, but the shape of the distribution is informative. The resultant curve produces a trough around 90°, consistent with the previous observer data. However, there is a marked peak in the red-magenta region at approx. 330°. A slight shoulder also occurs in the green hues around 175°.

The presence or absence of the red-magenta peak, and also a secondary green peak, in these datasets suggests a kind of duality depending on the viewing conditions, and which does not strictly adhere to traditionally defined object or aperture modes of viewing, particularly for simulated surface colors. It has been noted previously by Kuehni[21] that the H-K effect is rooted in the hue opponency of the human visual system. However, whilst the blue-yellow axis provides a peak and trough in the H-K effect common to all the datasets, the red-green axis provides a markedly separate peak in the red hues that does not appear in the object mode experiment.

4.2 | Extending the Fairchild and Pirrotta model

The Fairchild and Pirrotta 1991 model takes the form of Equations (3) and (4). The model predicts equivalent achromatic lightness L^*_{EAL} from a color's metric lightness L^* and chroma C^*_{ab}. The constants k_1 and k_2 apply a ½-cycle sine-based function, with a minimum effect at h_{ab} of 90° and maximum effect at 270°.

$$L^*_{\mathrm{EAL}} = L^* + f_{BY}(h^\circ) \cdot C^*_{ab} \tag{3}$$

$$f_{BY}(h^\circ) = k_1 \cdot \left| \sin\left(\frac{h^\circ - 90}{2}\right) \right| + k_2 \tag{4}$$

We therefore propose an extended model that takes into consideration the peak H-K effect in the red-magenta hues. The extended model takes a similar form, with an additional hue-dependent term (see Equation (5)). The previous term $f_{BY}(h^\circ)$ remains unchanged (see Equation (6)). The new term $f_R(h^\circ)$ is derived from a cosine-based function acting upon reddish hues across the range $h_{ab} \leq 90°$ or $\geq 270°$ with a broad distribution (see Equation (7)). For all other hue angles the value of $f_R(h^\circ)$ is constrained to zero.

$$L^*_{\mathrm{EAL}} = L^* + (f_{BY}(h^\circ) + f_R(h^\circ)) \cdot C^*_{ab} \tag{5}$$

$$f_{BY}(h^\circ) = k_1 \cdot \left| \sin\left(\frac{h^\circ - 90}{2}\right) \right| + k_2 \tag{6}$$

$$f_R(h^\circ) = \begin{cases} k_3 \cdot |\cos(h^\circ)| + k_4, & \text{for } -90° \leq h^\circ \leq 90°. \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

4.3 | Applying the models

The models are applied to two datasets. Firstly, we utilize Wyszecki's 1967 dataset[19] (ceramic tiles in object viewing mode) after transforming it from its native Illuminant C to a D50 white point using a one-step CAT16 chromatic adaptation transform, with degree of adaptation D equal to 1.0. Secondly, we combine the near-white substrate colors used previously by High and Green[15] (soft-proofed on a D50 display) with the present dataset.

Predicted lightnesses are compared against the observer data, with the RMS error calculated from the ΔL^* differences. Coefficients k_1 and k_2 in the ½-cycle sine-based model, and coefficients k_1, k_2, k_3 and k_4 in the extended model are optimized with the values shown in

TABLE 1 Fitting the 1967 ceramic tile and the 2021 & 2022 display-based datasets

	½-cycle sine-based model		Extended model			
Dataset	k_1	k_2	k_1	k_2	k_3	k_4
Wyszecki (1967)	0.1759	0.0627	0.1821	0.0550	−0.0084	0.0141
High et al. (2021 & 2022)	0.1825	0.0909	0.1644	0.0603	0.1307	0.0060

Coefficients k_1 and k_2 are optimized for Equation (4) in the ½-cycle sine-based model, and k_1, k_2, k_3 and k_4 are optimized for Equation (6) and (7) in the extended model.

TABLE 2 RMS errors

Dataset	½-cycle sine-based model RMS error	Extended model RMS error
Wyszecki (1967)	0.91	0.89
High et al. (2021 & 2022)	2.06	1.66

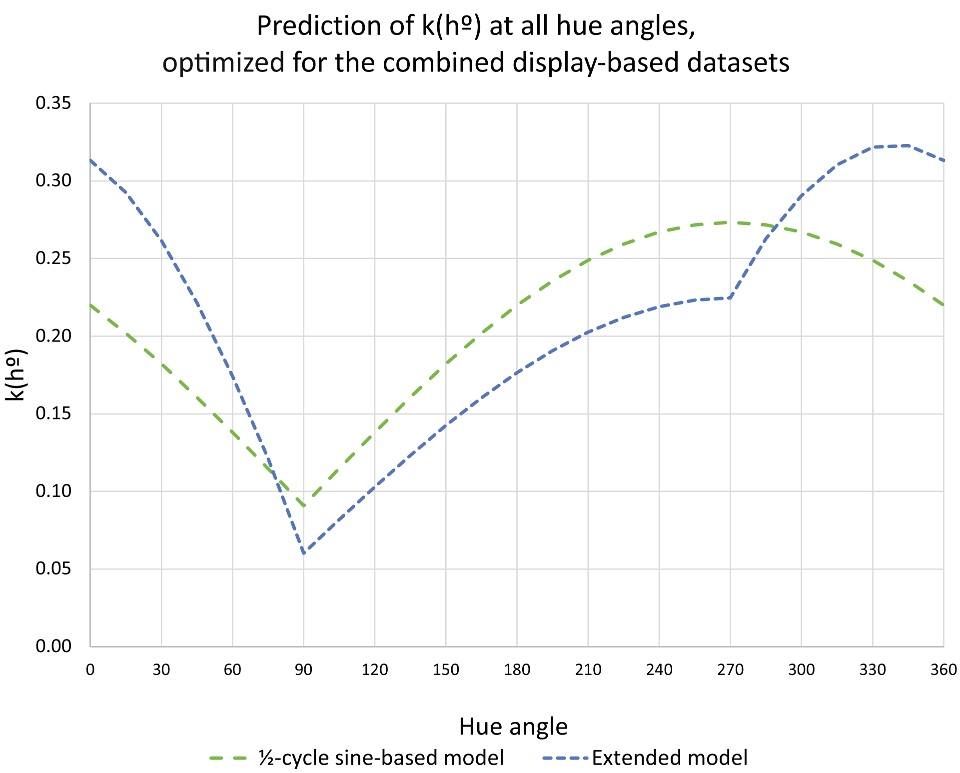

FIGURE 5 After fitting the combined display-based datasets, the two models (½-cycle sine-based and extended) are used to generate a value of $k_{(h^\circ)}$ at all hue angles

Table 1 to give the lowest RMS errors when applied to each of the two datasets. RMS errors are shown in Table 2.

4.4 | Effect of the extended model

When fitting the Wyszecki data we see that the additional term in the extended model has virtually a null effect when compared to the ½-cycle sine-based model. In Table 1 we see that the values for k_1 and k_2 remain almost unchanged, with k_3 and k_4 contributing very little. The RMS error therefore remains almost unchanged across the two models at 0.91 and 0.89 respectively (see Table 2).

By contrast, the extended model offers some improvement when fitting the combined 2021 and 2022 display-based data, with an RMS error of 1.66 compared to 2.06 for the ½-cycle sine-based model.

The source of this improvement can be illustrated by plotting the values of $k_{(h^\circ)}$ returned by the optimized models at each hue angle (see Figure 5).

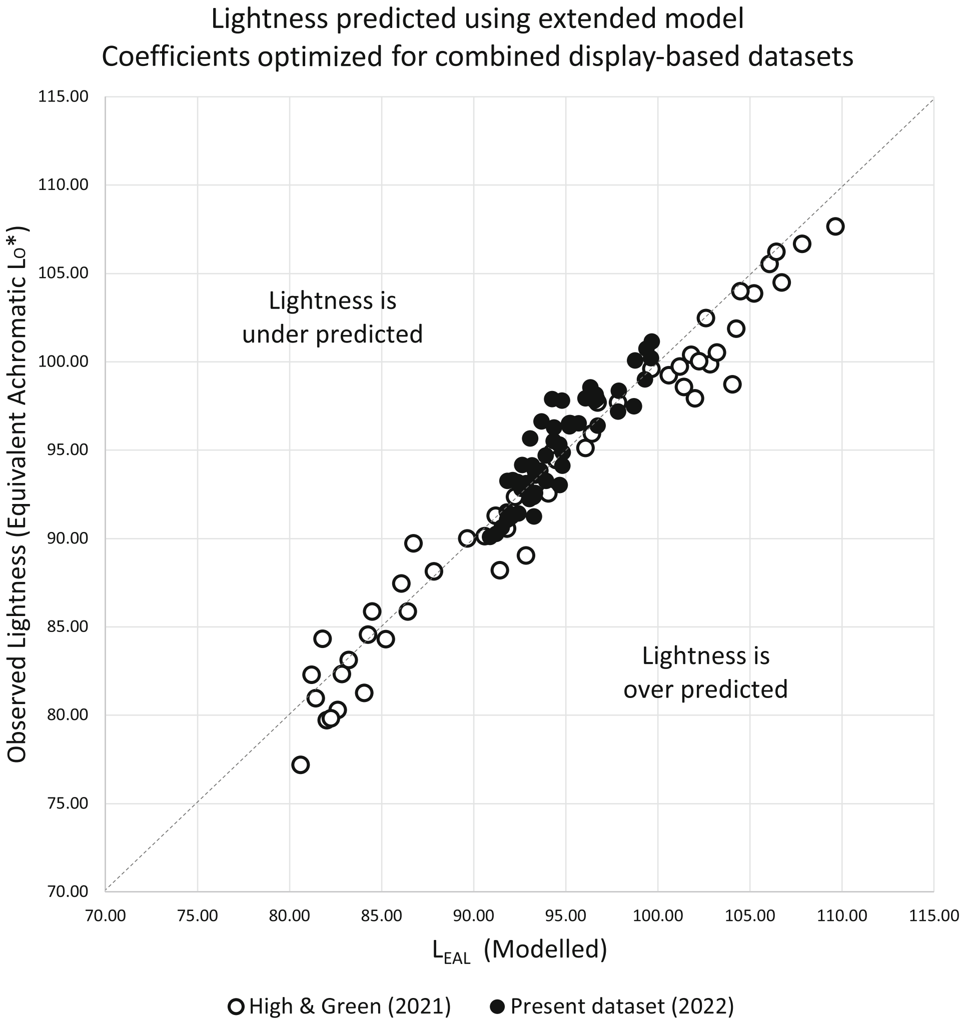

FIGURE 6 The extended model is optimized to fit the 2021 and 2022 datasets. The observed lightness (L^*_{O}) is plotted against the modelled equivalent achromatic lightness (L^*_{EAL}). The RMS error is 1.66

The ½-cycle sine-based model, when optimized for the display-based dataset, under-predicts the lightness of the red hues, whilst somewhat over-predicting the lightness of the green and blue hues. The extended model, however, is free to fit the lightness of the red hues independently of the green-blue hues. The predicted peak value of $k_{(h^\circ)}$ is approximately 0.32 in the magenta-red region, which is consistent with that directly derived from observer data in Figure 4.

4.5 | Performance of the extended model

Figure 6 plots the observed lightnesses (L^*_{O}) from the 2021 and 2022 display-based experiments against the modelled equivalent achromatic lightnesses (L^*_{EAL}) using the extended model.

Overall, Sanders and Wysecki's concept of constant L/Y ratio[5] is borne out, since the model predicts the perceived lightness over quite a wide range of luminances. This is true for lighter colors through to colors that would appear fluorescent or even self-luminous (with L^*_{O} of 80 through to 110).

The 2021 dataset from our previous experiment contains three groups of color stimuli with colorimetric L^* of 80, 90 and 100. These three groups are discernible in the plot as diagonal clusters which appear slightly off-axis. Closer inspection of the 2021 data reveals that it is generally lower chroma colors whose lightness is over-predicted. This detail is consistent with the previous findings[15] that the H-K effect is not readily observable in low chroma colors (C^*_{ab} of 10) when using the VAC method. However, when using the alternative VCC method in the 2021 experiment the H-K effect is indeed seen in lower chroma colors.

5 | DISCUSSION

By using a patch set with colors at every hue, it is possible to demonstrate the H-K effect as a continuous function (see Figure 3). Using only a small number of hues would be less time consuming, but this would mean that changes in other factors, such as viewing conditions and observer adaptation, could adversely affect the results without revealing their systematic effect.

The H-K effect was previously reported by Wyszecki as weak in the red-magenta region for mid-tone surface colors.[19] Although the Wyszecki's dataset does not specify inter-observer variability, it is clear that the data are well behaved, and this is reflected in the low RMS error returned by the modelled results. This may be due to the large number of observers, but also the mode of viewing. Comparative judgement between physical samples under a broadband light source, and against achromatic references set at narrow intervals of perceptual difference, provides very reliable results.

A method of adjustment is a practical alternative, but it has been seen to give higher inter-observer variability even when using physical samples,[1] and higher still when both reference and test colors are rendered on a display.[15,11] A weakness of the VAC approach is that observers give only a single response to each match. In reality there is an upper and lower threshold of perceived difference, and this can be demonstrated by making two separate lightness adjustments in an up and down direction.[11] A method of limits could also have provided these thresholds, requiring a binary decision to be made by the observer such as "lighter, or not lighter" and "darker, or not darker". However, both these approaches greatly increase the length of the experiments. Based on our previous results,[15] it may be wise to use these alternative approaches when judging the H-K effect in low-chroma colors.

6 | CONCLUSIONS

In a soft-proofing environment, where areas of lighter color such as paper colors may be the brightest on view, the H-K effect is very strong in the red-magenta hues. The magenta peak has also been reported in some works on unrelated colors (see Kim et al. in Hellwig et al.[13]). Conversely, in previous investigations colors viewed in the object mode did not exhibit the peak in red-magenta hues, nor the possible secondary peak in the green hues. A single hue angle dependent function may therefore be insufficient to describe the H-K effect in all circumstances.

We propose an extended model which better predicts the perceived lightness of those colors that lie towards the upper bounds of simulated surface colors, including substrate colors in a soft-proof. In this example, the H-K effect is modelled for simulated colors with an L^* of 80, and into the self-luminous range up to an L^*_O of 110. This approach may also work well for unrelated colors.

6.1 | Future work

Sanders and Wyszecki[20] discussed the confusion in appearance terminology when presenting self-luminous stimuli that might be viewed as surface colors relative to a brighter background. Nayatani referred to these as "pseudo-object colors".[18] The separate modes of viewing described for object, aperture or illuminant as part of Judd's classification of optical attributes became an ASTM standard in 1961, and the terminology has remained largely unchanged.[22] The CIE's International Lighting Vocabulary[7] has provided some additional terminology in terms of surface and aperture colors, luminous and non-luminous colors, as well as related and unrelated colors. However, the definitions are given in terms of appearance to the observer rather than the nature of the physical stimuli themselves.

In this way, the use of display systems is often ill-defined in terms of viewing mode and color appearance: displays may be used to simulate surface appearance, such as in a controlled display-to-print match, or they can create high luminances and highly saturated colors which can only be viewed as self-luminous. A clearer vocabulary for display-based viewing conditions would be welcome, and would help greatly in the "cross-over" between viewing modes. It would also help avoid ambiguity when modelling the H-K effect.

AUTHOR CONTRIBUTION

Gregory High: Conceptualization, Methodology, Software, Investigation, Visualization, Writing - Original Draft. **Phil Green:** Writing - Review & Editing. **Peter Nussbaum:** Writing - Review & Editing, Supervision.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the time, care and efforts of all the observers. Thanks goes to Aslak Tøn and Maximilian Czech for supervising some of the observers at the Colourlab.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article, or were previously published and are available from the cited sources.

ORCID

Gregory High https://orcid.org/0000-0001-5315-0772

REFERENCES

[1] Fairchild MD, Pirrotta E. Predicting the lightness of chromatic object colors using CIELAB. *Color Res Appl.* 1991;16(6):385-393.

[2] Parsons JH. *An Introduction to the Study of Colour Vision*. 2nd ed. Cambridge University Press; 1924.
[3] Fairchild MD. *Color Appearance Models*. 3rd ed. John Wiley & Sons; 2013.
[4] Wyszecki G, Stiles WS. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. 2nd ed. John Wiley & Sons; 1982.
[5] Sanders CL, Wyszecki G. Correlate for lightness in terms of CIE-Tristimulus values. Part I *JOSA*. 1957;47(5):398-404.
[6] Sanders CL, Wyszecki G. L/Y ratios in terms of CIE-chromaticity coordinates. *JOSA*. 1958;48(6):389-392.
[7] CIE S 017/E:2020 ILV: International Lighting Vocabulary, 2nd edition. International Commission on Illumination (CIE); 2020.
[8] Luo MR, Hunt RWG. The structure of the CIE 1997 colour appearance model (CIECAM97s). *Color Res Appl*. 1998;23(3): 138-146.
[9] Gao C, Li C, Shi K, Luo MR, Pointer MR. CAM20u: an extension of CAM16 for predicting the color appearance of unrelated colors. *Color Res Appl*. 2021;46(4):749-758.
[10] Huang WJ, Yang Y, Luo MR. Verification of the CAM15u colour appearance model and the QUGR glare model. *Light Res Technol*. 2017;51(1):24-36.
[11] Xie H, Fairchild MD. G0 revisited as equally bright reference boundary. *29th Color and Imaging Conference* 2021; 2021(1): 247-252.
[12] Park Y. Determining the color appearance of Helmholtz-Kohlrausch effect for self-emissive displays. *J Inform Display*. 2022;23(3):235-242.
[13] Hellwig L, Stolitzka D, Fairchild MD. Extending CIECAM02 and CAM16 for the Helmholtz-Kohlrausch effect. *Color Res Appl*. 2022;47(5):1096-1104.
[14] Hellwig L, Fairchild MD. Brightness, lightness, colorfulness, and chroma in CIECAM02 and CAM16. *Color Res Appl*. 2022; 47(5):1-13.
[15] High G, Green P. The impact of the Helmholtz-Kohlrausch effect on the appearance of near-white paper colours. *29th Color and Imaging Conference* 2021; 2021(1):241-246.
[16] ISO 3664:2009 Graphic technology and photography – viewing conditions. Int Org Standard. 2009.
[17] ISO 14861:2015 Graphic technology – requirements for colour soft proofing systems. Int Org Stand. 2015.
[18] Nayatani Y. Simple estimation methods for the Helmholtz–Kohlrausch effect. *Color Res Appl*. 1997;22(6):385-401.
[19] Wyszecki G. Correlate for lightness in terms of CIE chromaticity coordinates and luminous reflectance*. *JOSA*. 1967;57(2): 254-257.
[20] Sanders CL, Wyszecki G. Correlate for brightness in terms of CIE color matching data. CIE proceeding 15th session, Vienna, 1963 1964: 221–230.
[21] Kuehni RG. An opponent-color model for the Sanders-Wyszecki Helmholtz-Kohlrausch effect dataset. *Color Res Appl*. 2000;25(4):292-293.
[22] ASTM E284-17 Standard Terminology of Appearance. ASTM International; 2017.

AUTHOR BIOGRAPHIES

Gregory High is a PhD candidate at the Colour and Visual Computing Laboratory, NTNU, Norway. The topic of his PhD research project is “A model of consistent colour appearance”.

Phil Green is Professor of Colour Imaging at the Colour and Visual Computing Laboratory, NTNU, Norway.

Peter Nussbaum is an associate professor of colour imaging at the Colour and Visual Computing Laboratory, NTNU, Norway. Dr. Nussbaum received an MSc from the Colour & Imaging Institute, University of Derby, GB, in 2002 and completed his PhD degree in imaging science in 2011 from the University of Oslo, Norway.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: High G, Green P, Nussbaum P. The Helmholtz-Kohlrausch effect on display-based light colors and simulated substrate colors. *Color Res Appl*. 2023;48(2):167-177. doi:10.1002/col.22839