Psychological Research (2018) 82:896–914
https://doi.org/10.1007/s00426-017-0880-8

ORIGINAL ARTICLE

Color and emotion: effects of hue, saturation, and brightness

Lisa Wilms[1] · Daniel Oberfeld[1]

Received: 3 November 2016 / Accepted: 8 June 2017 / Published online: 13 June 2017

Abstract Previous studies on emotional effects of color often failed to control all the three perceptual dimensions of color: hue, saturation, and brightness. Here, we presented a three-dimensional space of chromatic colors by independently varying hue (blue, green, red), saturation (low, medium, high), and brightness (dark, medium, bright) in a factorial design. The 27 chromatic colors, plus 3 brightness-matched achromatic colors, were presented via an LED display. Participants ($N = 62$) viewed each color for 30 s and then rated their current emotional state (valence and arousal). Skin conductance and heart rate were measured continuously. The emotion ratings showed that saturated and bright colors were associated with higher arousal. The hue also had a significant effect on arousal, which increased from blue and green to red. The ratings of valence were the highest for saturated and bright colors, and also depended on the hue. Several interaction effects of the three color dimensions were observed for both arousal and valence. For instance, the valence ratings were higher for blue than for the remaining hues, but only for highly saturated colors. Saturated and bright colors caused significantly stronger skin conductance responses. Achromatic colors resulted in a short-term deceleration in the heart rate, while chromatic colors caused an acceleration. The results confirm that color stimuli have effects on the emotional state of the observer. These effects are not only determined by the hue of a color, as is often assumed, but by all the three color dimensions as well as their interactions.

Introduction

In the scientific literature as well as in applied domains like interior design and architecture, systematic effects of color on emotion are frequently assumed (e.g., Adams & Osgood, 1973; Greene, Bell, & Boyer, 1983; Heller, 2008; Meerwein, Rodeck, & Mahnke, 2007; Schaie, 1961; Schauss, 1985; Suk & Irtel, 2010; Valdez & Mehrabian, 1994; Wright & Rainwater, 1962). For instance, light consisting predominantly of longer wavelengths (red) was proposed to result in higher arousal than light with medium or short wavelengths (blue or green) (e.g., Jacobs & Hustmyer, 1974; Kaiser, 1984; Walters, Apter, & Svebak, 1982; Wilson, 1966). However, a color is specified not only by its hue (e.g., green, red, blue, or yellow), but also by two additional perceptual color dimensions: saturation (difference to an achromatic stimulus, i.e., a neutral gray or white) and brightness (perceived intensity of the light) (cf. Wyszecki & Stiles, 2000). Several studies indicated that the saturation of a color stimulus has a stronger effect on the emotional response than the hue (e.g., Gao et al., 2007; Suk & Irtel, 2010; Valdez & Mehrabian, 1994).

Unfortunately, many previous studies on emotional effects of color did not provide a sufficiently precise colorimetric specification of the stimuli or failed to control for saturation and brightness when presenting stimuli with different hues (see Elliot & Maier, 2014; Kaiser, 1984).

This experiment was conducted as a part of LW's master's thesis in Psychology (Institute of Psychology, Johannes Gutenberg-Universität Mainz, 2013). Portions of this work were presented at the Tagung experimentell arbeitender Psychologen 2014, Gießen, Germany, at the Conference of the International Society for Research on Emotion 2015, Geneva, Switzerland, and at the international symposium Seeing Colors 2016, Regensburg, Germany.

✉ Daniel Oberfeld
oberfeld@uni-mainz.de

[1] Institute of Psychology, Section Experimental Psychology, Johannes Gutenberg-Universität, 55099 Mainz, Germany

Several studies even used color words (e.g., "RED") rather than visual stimuli actually differing in color (e.g., Adams & Osgood, 1973). Here, we presented colorimetrically controlled stimuli defined in terms of the CIELAB color space (Commission Internationale de l'Éclairage, 2007), which is aimed at perceptual uniformity (Robertson, 1990), and we independently varied hue, saturation, and brightness in a factorial design.

A second shortcoming of many previous studies on effects of color on emotion was that associations between emotional terms (e.g., adjectives) and color were investigated, rather than measuring the actual emotional state induced by the color stimulus. For example, the study by Sato, Kajiwara, Hoshino, and Nakamura (2000) presented colorimetrically defined color stimuli, but the task of the participants was to "[…] select a word to describe that color from the set of 12 descriptive pairs of words […]" (p. 54), as for example "warm–cool". Thus, the participants rated the presented color on the semantic differentials, rather than rating their own current emotional state while viewing this color. Even though the emotional connotations of the colors (Soriano & Valenzuela, 2009; Wright & Rainwater, 1962) may influence the emotional state of the participant, this type of experiment does not provide direct information concerning emotional effects of color stimuli (cf. Whitfield & Wiltshire, 1990). Along the same line of reasoning, studies on color preferences (e.g., Guilford & Smith, 1959; Hurlbert & Ling, 2007; Ou, Luo, Woodcock, & Wright, 2004a; Palmer & Schloss, 2010) could be viewed as measuring the valence of colors. However, a statement such as 'I like this color.' is qualitatively different from the statement 'This color makes me feel good.' In the present study, we were interested in the latter, direct assessment of effects of color on emotion and explicitly instructed the participants to rate their current emotional state while viewing a color stimulus.

A second approach for gaining a better understanding of emotional responses to color is to complement measures of emotion experience (e.g., semantic differentials or rating scales) with physiological measures. Many physiological parameters have been used to study emotional processes (e.g., Ashcroft, Guimaraes, Wang, & Deakin, 1991; Bradley, Codispoti, Cuthbert, & Lang, 2001; Carvalho, Leite, Galdo-Alvarez, & Goncalves, 2012; Cuthbert, Schupp, Bradley, Birbaumer, & Lang, 2000; Gomez, Zimmermann, Guttormsen-Schar, & Danuser, 2005; Knutson, Katovich, & Suri, 2014; Lane et al., 1997; Lang et al., 1998; Palomba, Sarlo, Angrilli, Mini, & Stegagno, 2000; Roedema & Simons, 1999; Witvliet & Vrana, 1995). In particular, the activity of the autonomic nervous system is related to emotion (Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000; Kreibig, 2010; Mauss & Robinson, 2009). In the present study, we measured electrodermal (skin conductance response; SCR) and cardiovascular (heart rate) parameters, which were proposed as correlates of arousal (Lang, Greenwald, Bradley, & Hamm, 1993; Mauss & Robinson, 2009) and valence (Cacioppo et al., 2000), respectively.

Various authors report a positive correlation between emotional arousal and the amplitude of the SCR, for example when viewing emotional pictures (Bradley et al., 2001; Cuthbert et al., 2000; Lang et al., 1993; Roedema & Simons, 1999). SCR's latency is also shorter in high arousal conditions compared to low arousal conditions (Witvliet & Vrana, 1995). Furthermore, phasic and tonic skin conductance parameters tend to increase with the presented material's valence (Cuthbert et al., 2000; Gomez et al., 2005; Lang et al., 1993).

There is evidence for a positive correlation between the valence of pictures and the heart rate (Cacioppo et al., 2000; Lang et al., 1993). Concerning the relation between heart rate and arousal, results are somewhat disparate. Some authors report an increase of the heart rate with arousal (e.g., Lang et al., 1993; Witvliet & Vrana, 1995), some a decrease (e.g., Bradley et al., 2001), or no effect at all (e.g., Gomez et al., 2005).

Concerning physiological measures of color-induced emotion, research has focused on the effect of hue on arousal. For example, the skin conductance level was reported to be higher and the skin conductance response to be stronger when viewing red compared to blue or green stimuli (Gerard, 1958; Jacobs & Hustmyer, 1974; Nourse & Welch, 1971; Wilson, 1966), although a study by Mikellides (1990) reported no effects. Gerard (1958) and Jacobs and Hustmyer (1974) found no effects of hue on the heart rate, while Küller, Mikellides, and Janssens (2009) reported that the heart rate was higher when participants sat in a blue room compared to a red room. Comparable effects have been reported at the level of the central nervous system. Red light was reported to evoke higher cortical arousal measured by EEG than blue or green light (e.g., Ali, 1972; Carterette & Symmes, 1952). Furthermore, neuroimaging techniques show differences in the activation of hypothalamus and amygdala when blue or green light was presented (e.g., Vandewalle et al., 2010).

However, as mentioned above, most previous studies measuring physiological effects of color suffered from methodological problems concerning the control of the three color dimensions, hue, saturation, and brightness. In our study, we independently varied hue, saturation, and brightness in terms of the CIE LCh color space, which is the cylindrical representation of the CIE L*a*b* color space (Commission Internationale de l'Éclairage, 2007). We obtained ratings of the current emotional state of the participants while viewing a given color, and measured electrodermal and cardiovascular activity. Before the background of the previous results, we expected an effect of hue on arousal, with red eliciting higher arousal than blue even when controlling for brightness and saturation. We also expected a comparable or even stronger effect of saturation on arousal (Suk & Irtel, 2010; Valdez & Mehrabian, 1994). For valence, we expected the brightness to have the strongest effect (Valdez & Mehrabian, 1994). Because the SCR and the heart rate were reported to be associated with arousal and valence, respectively, we expected similar effects of brightness, saturation, and hue on these measures, and a correlation between the emotion ratings and the physiological measures.

Method

Participants

Sixty-five volunteer participants were recruited at the Johannes Gutenberg-Universität Mainz, Germany. Data from three participants had to be removed from the final analysis because of technical problems during the physiological recordings. Thus, 62 participants (49 women; age 19–54 years, mean age 23.37 years, SD = 6.0 years) remained in the analyses. All participants reported normal or corrected-to-normal visual acuity. They also had normal color vision, as confirmed by the Ishihara (2013) test for color blindness. The first 21 plates were shown under daylight conditions and if 17 or more plates were read correctly, the color vision was regarded as normal. Participants received partial course credit. The experiment was conducted in accordance with the principles expressed in the Declaration of Helsinki. The participants were naïve with respect to the hypotheses under test.

Apparatus and materials

The experiment was conducted in a completely darkened, electrically, and acoustically shielded double-walled cabin (Industrial Acoustics Company, Niederkrüchten) with a black interior lining. The room temperature was approximately 22 °C. Participants sat in an armchair positioned at a viewing distance of 1.5 m in front of an LED panel (Eurolite LED Panel RGB DMX) that presented the color stimuli. The dimensions of the display were 48 cm × 48 cm (height × width), corresponding to a visual angle of approximately 18.18° × 18.18°. The panel contains a total of 288 red, green, and blue LEDs, mounted under a transparent white diffuser. The LED panel was controlled by a PC and a DMX interface attached to the USB port. Instructions and rating scales were presented in white color on a black background on a 15″ computer screen located above the LED panel.

The skin conductance was measured using Ag/AgCl electrodes of 1 cm diameter, isotonic electrode gel, and a constant voltage of 0.5 V (Boucsein et al., 2012). The electrodes were affixed on the thenar and the hypothenar surfaces of the nondominant hand. The signals were amplified and converted to the digital domain at a sampling rate of 10 Hz (12 bit resolution). The electrocardiogram (ECG) was obtained from chest leads with Ag/AgCl electrodes of 1.5 cm diameter and isotonic electrode gel. One electrode was placed just below the right clavicle and the other one on the left side, on the lower lateral ribs. The ECG signals were sampled at a frequency of 512 Hz by the recording device Varioport (Kölner Vitaport-System, Becker Meditec). Skin conductance and ECG were both recorded continuously.

Color stimuli

Thirty different colors were presented in the experiment. In a 3 × 3 × 3 factorial design, three hues (blue, green, red) were combined with three saturation levels (low, medium, high) and three brightness levels (low, medium, high). In addition, participants were exposed to three achromatic stimuli (gray levels) on the same three brightness levels as for the chromatic stimuli. The stimuli were selected based on the Commission Internationale de l'Eclairage (CIE) 1976 *LCh* colorimetric system (i.e., the cylindrical representation of the CIE 1976 L*a*b system; Commission Internationale de l'Éclairage, 2007), in which each color is represented by the coordinates lightness (L^*), chroma (C^*), and hue (h^*) (Brainard, 2003; Wyszecki & Stiles, 2000). Due to the display's large visual angle, the 1964 CIE 10° standard observer (Commission Internationale de l'Éclairage, 2006) was used for computing the colorimetric values from the measured optical spectra (Photo Research PR-650 spectroradiometer). The lightness (L^*) is defined relative to a reference white and ranges from 0 to 100, where 100 represents the lightness of the reference white. We used the D65 standard illuminant as reference (Commission Internationale de l'Éclairage, 2008), with the CIE 1964 chromaticity coordinates (10° standard observer) $x_{\mathrm{ref}} = 0.31382$, $y_{\mathrm{ref}} = 0.33100$, and $Y_{\mathrm{ref}} = 223$ cd/m^2. Y_{ref} corresponds to the maximum luminance that can be produced by the LED panel. The chroma (C^*) measures "color purity" in terms of the distance from an achromatic stimulus (gray level) of the same brightness, for which $C^* = 0$. The hue angle ranges from $h^* = 0$ (red) through $\pi/2$ (yellow), π (green), $3/2\pi$ (blue), and back to 0. Because the chroma depends not only on "color purity" but also on brightness (Wyszecki & Stiles, 2000), and is, strictly speaking, not defined for stimuli viewed in isolation before a black background as in our study (Fairchild, 2005), the chroma was converted into saturation (sat), using the equation sat $= C^*/L^*$. Table 1 shows the colorimetric values of the presented stimuli. Table 2 shows the average hue angle, saturation, lightness, and luminance as a function of the three color dimensions that we varied in the experiment. The spectral peaks of the primaries were located at 625 nm (R), 515 nm (G), and 460 nm (B).

In the experiment, a neutrally gray adapting field with the values $L^* = 45.011$ ($Y = 32$ cd/m^2; CIE 1964 10° standard observer) and sat $= 0.0091$ was presented prior to each stimulus.

Response measures

Ratings of arousal and valence

In the study of emotion, a distinction is made between the discrete approach which considers different categories of basic emotions (e.g., Ekman, 1992) and the dimensional approach (e.g., Wundt, 1903) describing emotions with a small number of dimensions. In the latter approach, most authors propose three dimensions, often referred to as valence, arousal, and dominance (Mehrabian & Russell, 1974), or evaluation, activity, and potency (Osgood, Suci, & Tannenbaum, 1957). We used the dimensional approach.

Table 1 Colorimetric values of the 30 color stimuli

Hue	Brightness	Saturation	L^*	sat	h^*	Y (cd/m^2)	X	Z
Blue	High	High	50.00	2.31	5.06	40.43	56.74	324.06
		Medium	50.00	1.40	5.06	40.43	48.73	171.33
		Low	49.99	0.48	5.04	40.41	41.53	75.17
	Medium	High	34.99	2.28	5.07	18.64	25.24	126.98
		Medium	34.99	1.39	5.09	18.64	22.35	69.64
		Low	34.95	0.53	5.07	18.60	19.24	34.21
	Low	High	19.96	2.34	5.08	6.54	8.44	34.75
		Medium	19.89	1.45	4.99	6.50	7.17	21.20
		Low	20.10	0.52	5.06	6.62	6.72	11.04
Green	High	High	50.03	2.26	2.63	40.49	10.69	5.87
		Medium	50.00	1.39	2.65	40.42	18.48	15.74
		Low	50.01	0.50	2.61	40.45	30.30	30.40
	Medium	High	34.90	2.29	2.57	18.54	5.87	2.62
		Medium	34.99	1.46	2.66	18.65	8.86	7.76
		Low	34.99	0.56	2.63	18.65	13.90	14.19
	Low	High	20.19	2.33	2.48	6.67	2.78	0.92
		Medium	19.87	1.21	2.77	6.49	3.85	4.44
		Low	19.91	0.39	2.75	6.51	5.35	6.03
Red	High	High	50.00	2.31	0.79	40.43	81.42	0.64
		Medium	50.01	1.40	0.72	40.44	63.76	8.98
		Low	49.98	0.50	0.70	40.39	46.55	27.50
	Medium	High	34.98	2.35	0.78	18.63	35.91	0.42
		Medium	35.02	1.43	0.74	18.67	28.25	4.66
		Low	34.94	0.49	0.66	18.59	21.05	13.66
	Low	High	20.01	2.42	0.75	6.57	11.55	0.23
		Medium	19.94	1.31	0.41	6.53	9.54	4.02
		Low	20.24	0.43	0.47	6.69	7.34	5.91
Gray	Low	0	20.13	0.08		6.63	37.97	43.20
	Medium	0	35.01	0.05		18.67	17.36	20.74
	High	0	49.99	0.02		40.40	6.25	7.68

A factorial combination of hue (red, green, blue), saturation (low, medium, high), and brightness (low, medium, high) was presented, plus three equiluminant gray levels. The table displays the colorimetric values according to the CIE LCh (1976) system and a D65 reference white. The hue angle (h^*) is specified in radians. The saturation was defined as sat $= C^*/L^*$, where C^* and L^* is the chroma and lightness value, respectively, according to CIE LCh. The columns Y (luminance), X, and Z display the colorimetric coordinates according to the CIE 1964 system, 10° standard observer

Table 2 Average colorimetric values (CIE 1976 LCh system) for hue, saturation, lightness, and luminance (Y), averaged across the remaining dimensions for the set of 30 colors displayed in the experiment (see Table 1)

Color dimension	Mean	SD
Hue (h^*)		
Blue	5.058	0.03
Green	2.640	0.09
Red	0.668	0.14
Saturation (sat $= C^*/L^*$)		
Low	0.488	0.05
Medium	1.384	0.08
High	2.322	0.05
Achromatic	0.049	0.03
Lightness (L^*)		
Low	20.025	0.13
Medium	34.977	0.04
High	50.000	0.02
Y (cd/m^2)		
Low	6.57	0.07
Medium	18.63	0.04
High	40.43	0.03

For each color stimulus, the participants were asked to rate their current emotional state on the nonverbal "self assessment manikin" (SAM) scales for arousal and valence (Lang, 1980). These scales have proven to be useful for measuring emotional responses in various situations (Bradley & Lang, 1994) and have been used in connection with physiological measures (e.g., Bradley, Greenwald, Petry, & Lang, 1992; Gomez et al., 2005; Lang et al., 1993). The emotional dimensions valence and arousal are illustrated by five pictograms each. For the valence dimension, the scale ranges from a smiling, happy figure to a frowning, unhappy figure (see Fig. 2). For the arousal dimension, it ranges from a relaxed, sleepy figure to an excited, wide-eyed figure. A computer-based version of the valence and the arousal scale was used. In addition to the SAM pictograms, verbal descriptions of the emotional state were presented on each scale's left and right endpoints, based on semantic differentials from Mehrabian and Russell (1974) that show a high correlation with ratings on the SAM scales (Bradley & Lang, 1994). To the arousal scale, the German words "entspannt", "gelassen", and "träge" (relaxed, calm, and sluggish) were added on the left side (representing low arousal) and "angeregt", "aufgeregt", "erregt", "nervös", and "aufgerüttelt" (stimulated, agitated, excited, nervous, and startled) on the right side. The valence scale was presented with the words "unglücklich", "verärgert", "traurig", and "abgestoßen"(unhappy, angry, sad, and disgusted) on the left side and "angenehm", "fröhlich", and "vergnügt" (pleasant, cheerful, and joyful) on the right side. The scales were displayed with numerical labels ranging from "1" to "9", respectively. For the arousal scale, "1" corresponded to "calm" and "9" to "aroused". For the valence scale, "1" and "9" corresponded to "unpleasant" and "pleasant", respectively. The odd numbers were aligned vertically with the centers of the five pictograms, while the even numbers were placed exactly in the middle between two odd numbers. For each of the two scales, participants rated their current emotional state while viewing the color by entering one of the integer numbers 1–9 on a numerical keypad.

Electrophysiological measures

In addition to the assessment of arousal and valence with the SAM rating scales, several physiological parameters were measured. With respect to electrodermal activity (EDA), we analyzed the amplitude (SCR_{amp}) and latency (SCR_{lat}) of the skin conductance response (SCR). The SCR amplitude describes the difference between a response's base point and peak; the latency represents the time interval between stimulus onset and base point. SCRs were defined by a minimal amplitude of 0.02 μS and a latency between 1.0 and 3.5 s relative to stimulus onset (Boucsein et al., 2012). To identify SCRs from EDA recordings, we adapted a technique described by Storm, Fremming, Odegaard, Martinsen, and Morkrid (2000). First, the signal was smoothed by convolution with a Hanning window (length = 11 samples). Then, inflection points were determined to identify maxima and minima. The amplitude of a maximum (peak) was defined relative to the previous minimum (base point). According to our criteria concerning minimal SCR amplitude and time window, an SCR occurred only in 401 of the 1860 trials collected in the experiment (62 participants × 30 trials). When no SCR occurred, SCR_{amp} was set to the value 0 μS and SCR_{lat} was defined as missing. All skin conductance responses found by the computer algorithm were double-checked manually. For the analysis, the amplitudes were log-transformed ($\log_{10}[SCR_{amp} + 1]$) to make the distribution more symmetric (Boucsein et al., 2012).

As measures of cardiovascular activity, the stimulus-related change in heart rate was analyzed, which we term phasic heart rate (HR_{phasic}). Power line interference (50 Hz and harmonics) was removed by an FFT-based filter. In addition, a triangular filter was used to smooth the ECG. The heart beats were identified using an algorithm for R-wave detection (Vary & Stiny, 1977). Each ECG recording was entirely manually controlled and screened for signal dropouts and noise. For some participants, due to such technical problems, some sections of the ECG recordings could not be evaluated. These sections were excluded from the analyses by shortening the analysis time windows accordingly. In our analysis, phasic heart rate refers to a time window of 6 s after stimulus onset. The averaged time between consecutive heart beats (inter-beat interval, IBI) occurring in the time window of interest was determined and converted to beats per minute (bpm) by applying the formula HR (bpm) = 60000 (ms)/IBI (ms). The average HR in a time window of 6 s before stimulus onset was used as a baseline. The baseline HR value was subtracted from the HR value during the color stimulus, and the resulting stimulus-related change in HR was analyzed.

Design and procedure

In a within-subjects design, each of the 30 color stimuli was presented to each participant. To minimize order effects, the order of trials was randomized. Participants were tested individually. After providing written informed consent, participants performed the Ishihara (2013) test of color vision. Then, the electrodes were attached and general instructions about how to avoid possible disturbances of the skin conductance and ECG measurements and how to use the SAM rating scales were provided by the experimenter. During the experiment, detailed instructions were given via a computer screen positioned above the LED panel. The experiment started with three practice trials presenting colors not included in the main experiment. After that, 30 trials each containing one of the color stimuli were presented in randomized order. To control for the chromatic adaptation of rods and cones (Jameson, Hurvich, & Varner, 1979), each trial started with the presentation of a neutral white adaptation field on the LED panel for 60 s (see Fig. 1). During the initial 30 s of the trial, the participants performed a visual task to ensure that they kept looking at the adaptation field. Four white LEDs were positioned at the four corners of the LED panel. Each of the lamps was controlled independently and switched on for 500 ms with randomly varying inter-onset intervals (IOIs). For each lamp, the IOIs were drawn from an exponential distribution with $\lambda = 0.1$. With this distribution, the hazard function is constant, so that the time, since the last occurrence provides no information about the occurrence of the next event (Luce, 1991). The IOIs were limited to a minimum of 750 ms. Participants were instructed to count the total number of flashes for all four lamps. After 30 s, they were asked to enter the number via the computer keyboard.

To stabilize the electrodermal and cardiovascular activity before the onset of the color stimulus, the participants were instructed to sit quietly and look steadily at the center of the LED panel for the next 30 s (i.e., the second half of the 60 s adaptation field). Then, the color stimulus was switched on (60 s after trial onset) and participants were instructed to keep fixating the center of the LED panel. Thirty seconds after color stimulus onset (i.e., 90 s after trial onset), the SAM rating scales were presented on the computer screen, while the LED panel kept presenting the color stimulus. The participant first entered the arousal rating on the computer keyboard, followed by the valence rating. Immediately after the second SAM rating had been entered, the color stimulus was switched off and the next trial started with the adapting field.

Note that the viewing time of the color stimulus varied slightly, because we did not impose a response deadline for the SAM ratings. However, the variation in viewing time was relatively small. For the first SAM rating (arousal), the mean RT (relative to color stimulus onset) was 37.6 s

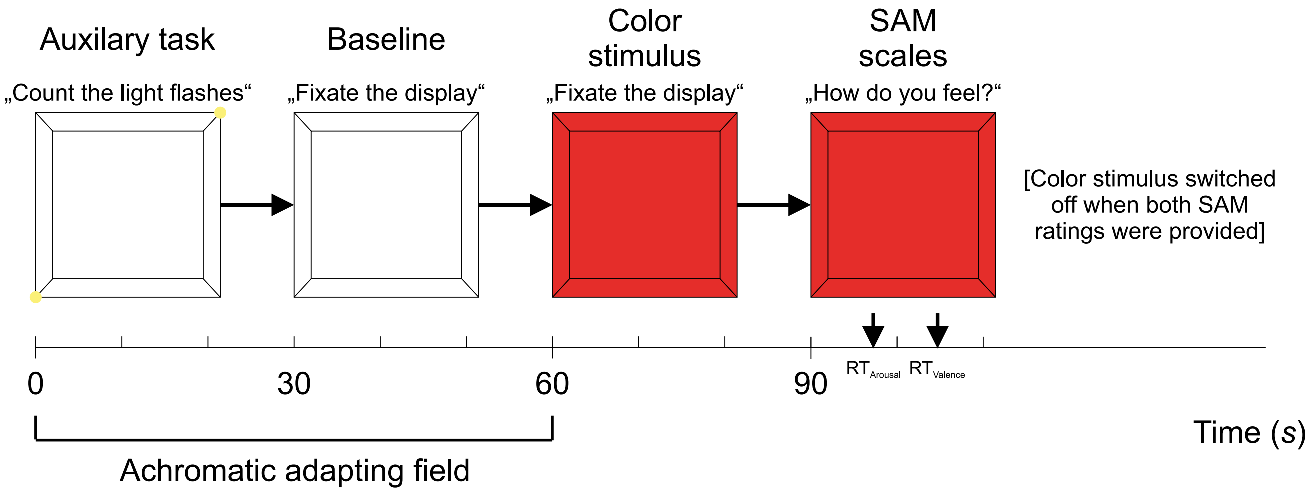

Fig. 1 Schematic description of the temporal structure of a trial. $RT_{Arousal}$: example response time for the arousal rating, $RT_{Valence}$: example response time for the valence rating

(SD = 4.6 s). 80% of the RTs were in the interval between 33.4 and 42.6 s. For the second SAM rating (valence), the RT variability was somewhat higher, because due to the sequential collection of first the arousal rating and then the valence rating, the variation in the arousal rating RT contributed to the variation in the valence rating RT. The mean valence rating RT (relative to color stimulus onset) was 45.0 s (SD = 7.2 s). 80% of the RTs were in the interval between 37.2 and 53.6 s. It is unlikely that these relatively small variations in viewing time caused a substantial variability in chromatic adaptation.

Data analysis

For all chromatic colors, means of each dependent variable were analyzed separately in a 3 × 3 × 3 repeated-measures ANOVA using a multivariate approach (Rouanet & Lépine, 1970). The within-subjects factors were hue, saturation, and brightness. Partial η^2 is reported as a measure of effect size. With regard to the analysis of $\mathrm{SCR_{lat}}$, only 401 (out of 1860) trials could be analyzed, as discussed above. As a consequence, there were missing values for several participants. For this reason, a maximum-likelihood procedure (SAS PROC MIXED) was used for analyzing $\mathrm{SCR_{lat}}$. This procedure is suitable for data with missing values (Keselman, Algina, & Kowalchuk, 2002). The degrees of freedom were computed by the method of Kenward and Roger (1997), which has been demonstrated to be superior to alternative methods (Arnau, Bono, & Vallejo, 2009; Keselman et al., 2002; Kowalchuk, Keselman, Algina, & Wolfinger, 2004).

To include the gray levels, two additional rmANOVAs were calculated for each dependent variable, again using the multivariate approach. In the first ANOVA, the arousal ratings were averaged across the four hues, and the within-subjects factors were saturation (low, medium, high, achromatic) and brightness (low, medium, high). In the second ANOVA, the arousal ratings were averaged across saturation, and the within-subjects factors were hue (blue, green, red, gray) and brightness (low, medium, high).

Results

Range of emotion ratings across the presented color space

Figure 2 shows the mean ratings of arousal and valence on the 9-point SAM rating scales for each of the 30 colors. The self-rated arousal while viewing the color stimuli showed a considerable variation, only slightly smaller than the range of arousal ratings obtained for the widely used International Affective Picture System (IAPS; Lang, Öhman, & Vaitl, 1998). However, all valence ratings were rather close to the neutral rating category (5), and thus showed considerably less variation than for the IAPS, where valence ratings range from 1 to more than 8 (see Fig. 1 in Lang et al., 1993).

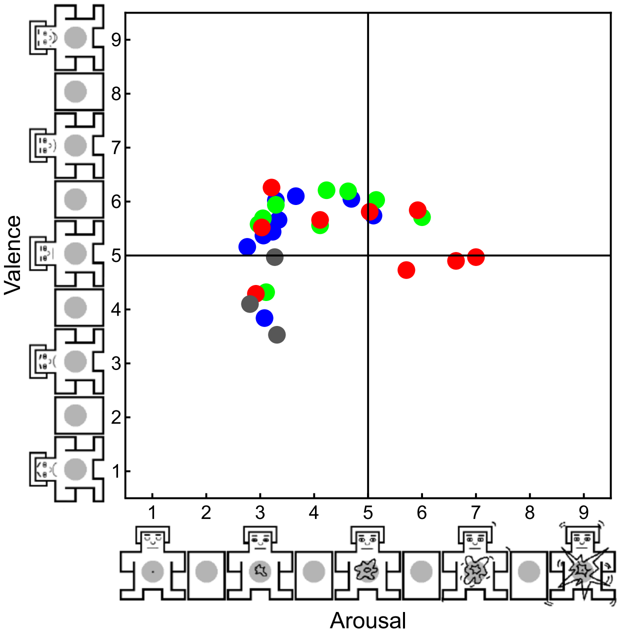

Fig. 2 Mean ratings of valence and arousal on the SAM rating scales for each of the 30 color stimuli. *Colors* indicate the hue of the corresponding color stimulus (*blue*, *green*, *red*, and *gray*) (color figure online)

Ratings of arousal

For the chromatic colors (blue, green, and red), a three-factorial rmANOVA with the within-subjects factors hue, saturation, and brightness showed a significant effect of hue, $F(2, 60) = 42.840$, $p < 0.001$, $\eta_p^2 = 0.588$. Panel a in Fig. 3 illustrates that for the high and medium saturation levels, arousal ratings increased from blue to green and from green to red. As a post hoc analysis, pairwise comparisons were conducted via paired-samples t tests. Here and in the following, the Hochberg (1988) method was used to control the family wise error rate. At an α-level of 0.05, all the three pairwise comparisons between the mean arousal ratings at the three chromatic hues were significant. The effect sizes in terms of d_z (Cohen, 1988) ranged from 0.59 to 1.18. Cohen (1988) classified values between 0.2 and 0.5 as small, values between 0.5 and 0.8 as medium, and values greater than 0.8 as large.

Saturation also showed a significant effect on arousal, $F(2, 60) = 67.747$, $p < 0.001$, $\eta_p^2 = 0.693$. Panels a, c in Fig. 3 show that the average arousal ratings were higher for highly saturated chromatic colors than for colors with medium or low saturation. Post hoc pairwise comparisons

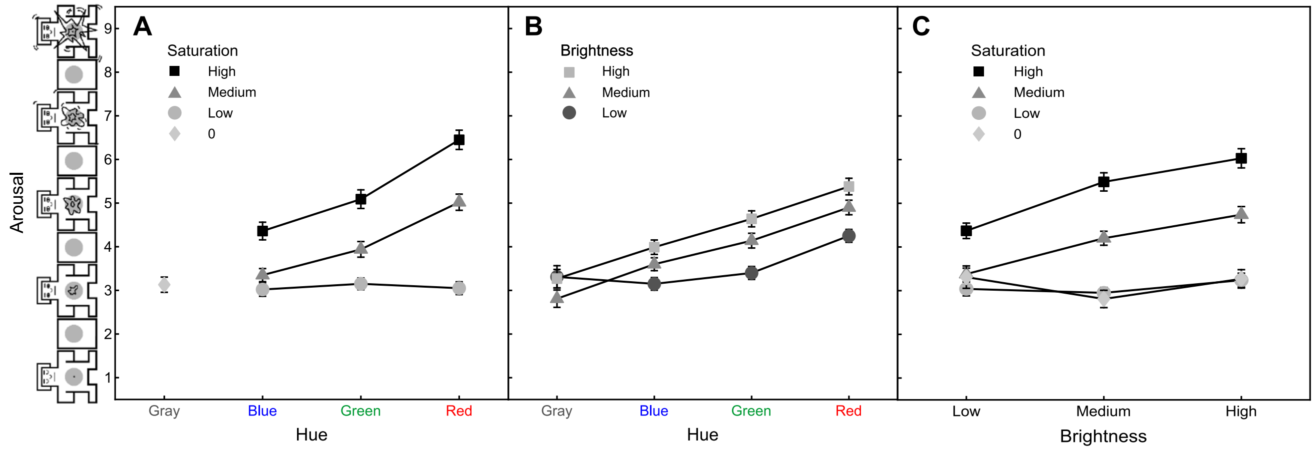

Fig. 3 Average arousal ratings. **a** As a function of hue and saturation. *Boxes* high saturation, *triangles* medium saturation, *circles* low saturation, *diamonds* saturation 0 (gray levels). **b** As function of hue and brightness. *Boxes* high brightness, *triangles* medium brightness, *circles* low brightness. **c** As a function of brightness and saturation. *Error bars* show ±1 standard error of the mean (SEM) across the 62 participants

showed significant differences in self-rated arousal between all the three pairs of saturations, with effect sizes ranging from $d_z = 1.01$ to 1.49.

There was also a significant effect of brightness on arousal, $F(2, 60) = 25.461$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.459$. As shown in Fig. 3c, averaged across saturation, the arousal ratings increased with brightness. Post hoc pairwise comparisons showed significant differences in self-rated arousal between all of the three pairs of brightness, with effect sizes ranging from $d_z = 0.66$ to 0.91. Figure 3a shows that the increase of the arousal ratings from blue to green to red was only present at the two higher saturation levels. The hue × saturation interaction was significant, $F(4, 58) = 25.417$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.637$. The brightness × hue interaction was not significant, $F(4, 58) = 1.350$, $p = 0.263$, $\eta_{\mathrm{p}}^2 = 0.085$ (see Fig. 3b), but the hue × saturation × brightness interaction was significant, $F(8, 54) = 2.657$, $\mathrm{p} = 0.016$, $\eta_{\mathrm{p}}^2 = 0.282$. Figure 3c shows an interaction of brightness and saturation, $F(4, 58) = 17.259$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.543$, the increase of the arousal ratings from low to medium to high brightness was only present at the two higher saturation levels.

Figure 3a shows that the arousal ratings for the gray levels were similar to the ratings for the chromatic stimuli with low saturation. Figure 3b demonstrates that the arousal ratings while viewing the gray levels did not change strongly across the brightness levels. To include the gray levels, two additional two-factorial rmANOVAs were conducted on arousal. In the first ANOVA, the arousal ratings were averaged across hue, and the within-subjects factors were brightness and saturation (see Fig. 3c). There was a significant effect of brightness, $F(2, 60) = 14.06$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.32$, a significant effect of saturation, $F(3, 59) = 44.471$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.693$, and a significant brightness × saturation interaction, $F(6, 56) = 12.63$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.575$.

In the second ANOVA, the arousal ratings were averaged across saturation, and the within-subjects factors were brightness and hue. The mean ratings are shown in Fig. 3b. The effect of brightness was of course identical to the ANOVA above. The effect of hue, $F(3, 59) = 34.396$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.636$, and the brightness × hue interaction, $F(6, 56) = 3.81$, $p = 0.003$, $\eta_{\mathrm{p}}^2 = 0.290$, were significant.

Taken together, all of the three color dimensions (hue, saturation, and brightness) clearly influenced the ratings of arousal while viewing the color stimulus.

Ratings of valence

A repeated-measures ANOVA including all chromatic colors with the within-subjects factors brightness, saturation, and hue showed a significant effect of saturation, $F(2, 60) = 15.627$, $p < 0.001$, $\eta_{\mathrm{p}}^2 = 0.343$. On average, the valence rating was highest (i.e., most positive) for the chromatic colors with medium saturation ($M = 5.82$, $\mathrm{SD} = 1.64$), followed by the colors with high and then low saturation ($M = 5.52$, $\mathrm{SD} = 1.81$ and $M = 5.16$, $\mathrm{SD} = 1.77$). Post hoc pairwise comparisons showed significant differences in self-rated valence between all of the three pairs of saturation, with effect sizes ranging from $d_z = 0.26$ to 0.69.

Panel a in Fig. 4 shows that on average, the most positive ratings came along with green hue ($M = 5.69$, $\mathrm{SD} = 1.69$), followed by blue ($M = 5.49$, $\mathrm{SD} = 1.79$) and then red ($M = 5.33$, $\mathrm{SD} = 1.80$). However, the effect of hue just failed to reach significance, $F(2, 60) = 3.128$, $p = 0.051$, $\eta_{\mathrm{p}}^2 = 0.094$. The observed pattern deviates

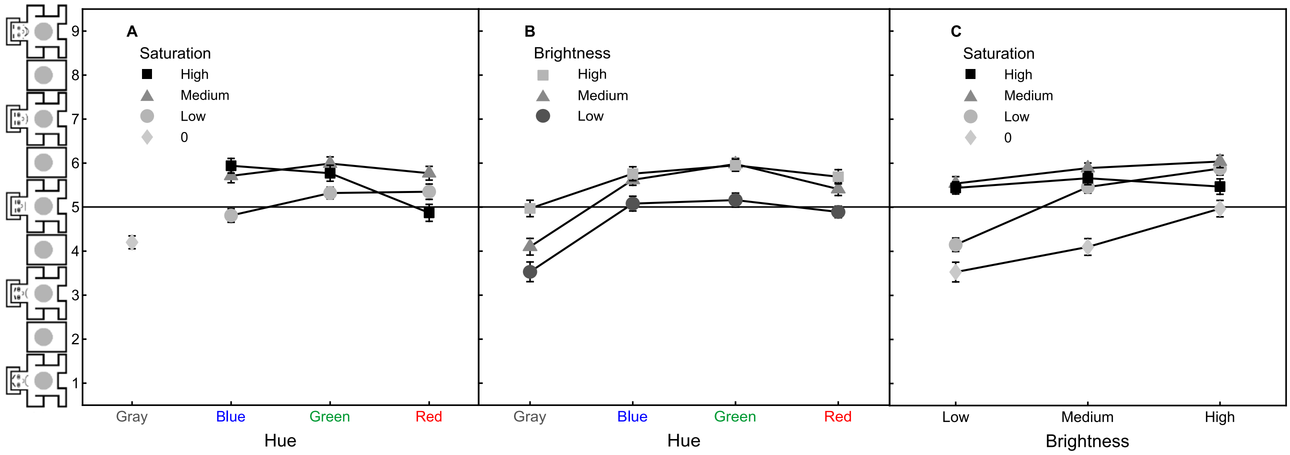

Fig. 4 Average valence ratings. **a** As a function of hue and saturation. **b** As a function of hue and brightness. **c** As a function of brightness and saturation. The same format as Fig. 3

slightly from the previous results (Oberfeld, Hecht, Allendorf, & Wickelmaier, 2009; Suk & Irtel, 2010; Valdez & Mehrabian, 1994), and also from the color preference studies mentioned earlier (e.g., Guilford & Smith, 1959; Terwogt & Hoeksma, 1995), which reported blue as the favored or most pleasant color. However, in our data, the valence ratings for blue and green differed only slightly. Figure 4a shows a different dependence of the valence ratings on the hue at the different saturation levels. Only for colors with high saturation, the typical color preference ranking blue–green–red was found. At medium saturation, the valence ratings were approximately equal across the three chromatic hues, while at low saturation, the highest valence rating was observed for red, showing the opposite pattern as compared to high saturation. This saturation × hue interaction was significant, $F(4, 58) = 11.223$, $p < 0.001$, $\eta_p^2 = 0.436$. Three separate post hoc rmANOVAs conducted at each saturation level showed a significant effect of hue (Hochberg-corrected) at low and high saturation, but not at medium saturation.

There was a significant effect of brightness on valence, $F(2, 60) = 28.933$, $p < 0.001$, $\eta_p^2 = 0.491$. As can be seen in Fig. 4b, the average valence ratings increased with brightness. Post hoc pairwise comparisons showed significant differences in self-rated valence between low and medium as well as between low and high brightness ($d_z = 0.77$ and 0.97, respectively), but not between medium and high brightness.

The brightness × hue interaction was not significant, $F(4, 58) = 1.329$, $p = 0.270$, $\eta_p^2 = 0.084$. However, the brightness × saturation interaction was significant, $F(4, 58) = 11.987$, $p < 0.001$, $\eta_p^2 = 0.453$. Figure 4c displays that valence ratings for stimuli with high saturation depended only weakly on brightness, while a strong increase in valence from low to medium to high brightness was observed for stimuli of low saturation. The brightness × saturation × hue interaction just failed to reach significance, $F(8, 54) = 2.118$, $p = 0.050$, $\eta_p^2 = 0.239$.

Including both the chromatic and achromatic stimuli, two separate rmANOVAs were conducted on the valence ratings, averaging across either hue or saturation. In the ANOVA with the within-subjects factors brightness and saturation, brightness, $F(2, 60) = 29.9$, $p < 0.001$, $\eta_p^2 = 0.50$, and saturation, $F(3, 59) = 30.527$, $p < 0.001$, $\eta_p^2 = 0.608$, had a significant effect on the valence ratings. As shown in Fig. 4c, the gray levels were rated as being less pleasant than the chromatic stimuli. The brightness × saturation interaction was also significant, $F(6, 56) = 8.52$, $p < 0.001$, $\eta_p^2 = 0.48$.

In the ANOVA with the within-subject factors, brightness and hue, there was a significant main effect of hue on valence, $F(3, 59) = 34.000$, $p < 0.001$, $\eta_p^2 = 0.634$, reflecting predominantly the difference in valence between chromatic and achromatic stimuli. The brightness × hue interaction was significant, $F(6, 56) = 2.57$, $p = 0.029$, $\eta_p^2 = 0.22$ (see Fig. 4b).

To summarize, although the variation in the valence ratings across the 30 colors was smaller than for saturation (see Fig. 2), all the three perceptual color dimensions had an effect. The valence mainly depended on whether a chromatic color or a gray level was presented, but several other significant main effects and interactions were also observed.

Arousal and valence ratings: sex differences

To answer the question whether the arousal and valence ratings differed between females and males, we conducted additional rmANOVAs with the between-subjects factor sex (female, male). The majority (79%) of our participants were women, as it is typical for samples of Psychology students in Germany. Because the traditional procedures for rmANOVAs require the design to be balanced (cf. Keselman, Algina, & Kowalchuk, 2001), the analysis was performed in R 3.3.3 (R Core Team, 2017), using the HRM package (Happ, Harrar, & Bathke, 2017) with the functions *hrm_test* and *hrm.test.dataframe*. The theoretical background and a detailed description of this R package can be found in Happ, Harrar, and Bathke (2016) and Happ et al. (2017). The procedure shows sufficient control of the Type I error rate even when the design is unbalanced and the variance–covariance matrices differ between groups.

All of the rmANOVAs on arousal and valence ratings reported above were repeated with the additional between-subjects factor sex. There was a significant hue × sex interaction in the analysis of the valence ratings including achromatic colors and averaging across saturation, $F(3.09, 68.57) = 3.79$, $p = 0.013$. The average valence rating for gray colors was lower in females ($M = 3.99$, $SD = 1.04$) than in males ($M = 4.97$, $SD = 1.25$), $t(16.66) = 2.60$, $p = 0.019$ (two-tailed). For the three chromatic hues, the average valence ratings were very similar between the two sexes. In the analysis of the valence ratings including achromatic colors and averaging across saturation, the saturation × sex interaction was significant, $F(2.51, 67.07) = 4.89$, $p = 0.006$. Note that this analysis includes exactly the same pair of average valence ratings for gray colors (i.e., 0 saturation) as above. Similar to the analysis above, the valence ratings for the remaining saturations were very similar between the two sexes. In all other rmANOVAs, the main effect of sex and all interaction effects involving sex were not significant (all p values >0.168).

Amplitude of skin conductance responses (SCR_{amp})

The averaged (log-transformed) amplitudes of the skin conductance responses for all chromatic colors were analyzed via an rmANOVA with the within-subjects factors, hue, saturation, and brightness. One should note that for 12 of the 62 participants, no SCRs occurred during the 30 trials (according to the criteria specified in the methods section). For all trials on which no SCR occurred, the SCR amplitude was set to 0 µS. We did not exclude these participants from the analysis, which of course attenuates the effects. There was a significant effect of saturation, $F(2, 60) = 6.75$, $p = 0.002$, $\eta_p^2 = 0.184$. As can be seen in Fig. 5, the highest mean amplitudes were observed at the high saturation level, followed by the medium and then by the low saturation level. The pattern shown in Fig. 5 is similar to the pattern of arousal ratings shown in Fig. 3a. The effect of brightness was also significant, $F(2, 60) = 3.88$, $p = 0.026$, $\eta_p^2 = 0.114$. The log SCR amplitude increased with increasing brightness (low brightness: $M = 0.0246$, $SD = 0.035$; medium brightness: $M = 0.0304$, $SD = 0.0483$; high brightness: $M = 0.0346$, $SD = 0.0526$), similar to the arousal ratings shown in Fig. 3b. There was no significant effect of hue, $F(2, 60) = 2.26$, $p = 0.113$, $\eta_p^2 = 0.070$, although descriptively, the SCR amplitude increased from blue to green to red, as in the arousal ratings. The saturation × hue interaction was not significant, $F(4, 58) = 0.96$, $p = 0.43$. There was a significant hue × saturation × brightness interaction, $F(8, 54) = 2.35$, $p = 0.031$, $= 0.258$. No other interaction effects were significant (all p values >0.270).

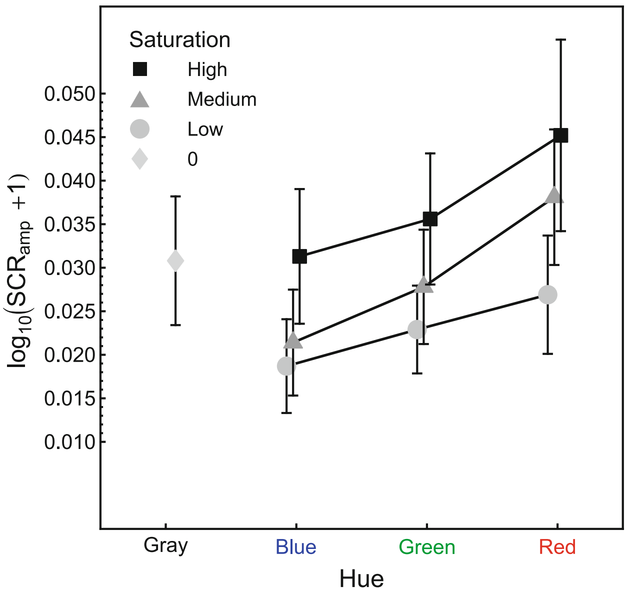

Fig. 5 Average log-transformed amplitude of the skin conductance response ($\log_{10}[SCR_{amp} + 1]$) as a function of hue and saturation. *Boxes* high saturation, *triangles* medium saturation, *circles* low saturation, *diamonds* saturation 0 (gray levels). *Error bars* show ±1 SEM

Figure 5 shows that the SCR amplitudes in response to gray stimuli were comparable to the mean amplitudes to chromatic colors. For all color stimuli (including gray), an rmANOVA with the within-subjects factors brightness and saturation and an rmANOVA with the within-subjects factors brightness and hue were conducted. In the former analysis, there was a trend towards a significant effect of saturation, $F(2, 60) = 2.62$, $p = 0.059$, $\eta_p^2 = 0.118$, and towards a significant effect of brightness, $F(2, 60) = 2.58$, $p = 0.084$, $\eta_p^2 = 0.079$. In the latter analysis, the effect of hue was significant, $F(3, 59) = 4.52$, $p = 0.006$, $\eta_p^2 = 0.187$, which can be attributed to the higher average SCR amplitude observed with achromatic compared to chromatic stimuli. The interactions between brightness and saturation and between hue and brightness were not significant (both p values >0.33).

Latency of skin conductance responses (SCR_{lat})

As mentioned above, when no SCR occurred according to the criteria defined in the Methods section, the latency was defined as missing. Therefore, there were missing values for some combinations of participant and color. For this reason, the effects of hue, saturation, and brightness on the SCR latency were analyzed with a maximum-likelihood procedure (SAS PROC MIXED) with Kenward and Roger (1997) adjusted degrees of freedom. The model covariance structure was autoregressive-heterogeneous (ARH(1); Wolfinger, 1996). The analysis showed no significant effects of hue, saturation, or brightness on SCR_{lat}, nor any significant interaction effects (all p values >0.28).

It remains unclear whether the non-significant effects of color on SCR_{lat} can be attributed to the small data basis due to missing values, or whether this indicates that the SCR latency is not related to affective responses to color presentations.

Phasic heart rate (HR_{phasic})

For computing the phasic heart rate, the average HR in a time window of 6 s before color stimulus onset was used as a baseline and subtracted from the average HR in a time window of 6 s starting at stimulus onset. For the chromatic colors, HR_{phasic} was analyzed via an rmANOVA with the within-subjects factors hue, saturation, and brightness. The average values of HR_{phasic} are displayed in Fig. 6. There was a trend towards an effect of brightness, $F(2, 60) = 2.57$, $p = 0.085$, $\eta_p^2 = 0.079$. Descriptively, the average heart rate acceleration was higher at the highest brightness level ($M = 0.355$, $SD = 1.60$), than at the intermediate ($M = 0.287$, $SD = 1.47$) or low ($M = 0.297$, $SD = 1.57$) brightness level. All other main effects and interaction effects were non-significant (all p values >0.159).

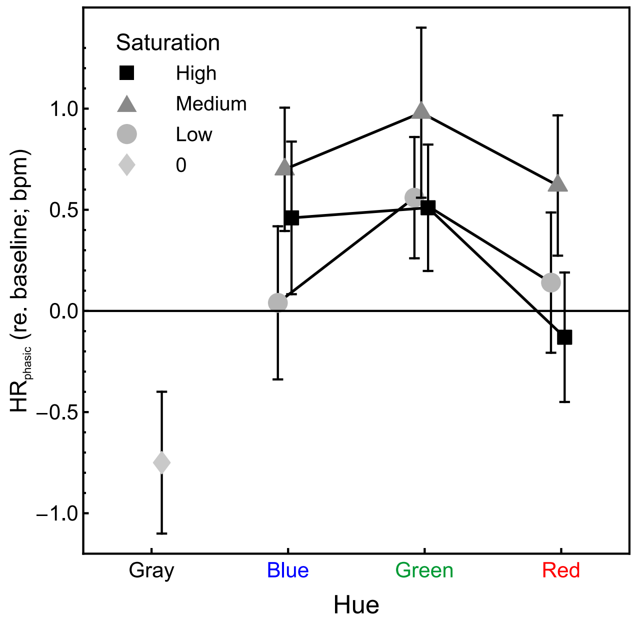

Fig. 6 Average phasic heart rate (HR_{phasic}) as a function of hue and saturation. *Boxes* high saturation, *circles* medium saturation, *triangles* low saturation, *diamonds* saturation 0 (gray levels). *Error bars* show ±1 standard error of the mean (SEM)

Figure 6 shows that gray stimuli caused a short-term heart rate deceleration, while chromatic stimuli resulted in a small acceleration. To include the achromatic stimuli, two additional ANOVAs were conducted, one with the within-subjects factors brightness and saturation, and the other with the within-subjects factors brightness and hue. The former ANOVA showed a significant effect of saturation, $F(3, 59) = 6.253$, $p = 0.001$, $\eta_p^2 = 0.241$, and the latter ANOVA showed a significant effect of hue, $F(3, 59) = 5.636$, $p = 0.002$, $\eta_p^2 = 0.223$. Figure 6 suggests that both effects can be attributed to a difference in the effects of achromatic and chromatic stimuli on HR_{phasic}. All other effects were non-significant (all p values >0.236).

Correlations between emotion ratings and physiological measures

To analyze the relation between the dependent variables, we used a subject-specific random-effects linear regression model (cf. Burton, Gurrin, & Sly, 1998), taking into account the repeated-measures structure of the data. For each participant, the emotion ratings and the physiological parameters were first z-standardized, and we used regression through the origin with a random slope. In this approach, the slope of the regression line relating, e.g., the arousal rating and the SCR amplitude across the 30 colors within a given participant, is assumed to vary from participant to participant, and the correlation structure is modeled by treating the participants as a random sample from the population of all possible participants. The variance–covariance matrix was specified as being of type "unstructured" (UN; Wolfinger, 1996), and the degrees of freedom were computed according to the method by Kenward and Roger (1997). We computed the coefficient of determination for fixed effects according to Edwards, Muller, Wolfinger, Qaqish, and Schabenberger (2008). Rather than analyzing whether an individual who provides, e.g., a high average arousal rating also shows a large average SCR amplitude (cf. Bland & Altman, 1995b), this within-subjects correlation (cf. Bland & Altman, 1995a) quantifies the extent to which, across the 30 colors presented to each participant, an increase in for example the arousal rating was associated with an increase in the amplitude of the SCR. Of the 62 participants, 12 did not show any SCR across the 30 colors, according to the

Table 3 Correlations between the dependent variables (z-standardized per participant)

	SAM arousal	SAM valence	log $\mathrm{SCR}_{\mathrm{amp}}$ ($N = 50$)	$\mathrm{SCR}_{\mathrm{lat}}$ ($N = 41$)
SAM valence	**$\beta = -0.1447$** **$p = 0.0057$** **$R^2_\beta = 0.1188$**			
log $\mathrm{SCR}_{\mathrm{amp}}$ ($N = 50$)	**$\beta = 0.074$** **$p = 0.0040$** **$R^2_\beta = 0.0055$**	$\beta = 0.018$ $p = 0.545$ $R^2_\beta = 0.0075$		
$\mathrm{SCR}_{\mathrm{lat}}$ ($N = 41$)	$\beta = 0.002$ $p = 0.97$ $R^2_\beta = 0.000$	$\beta = 0.032$ $p = 0.598$ $R^2_\beta = 0.0116$	**$\beta = -0.336$** **$p < 0.001$** **$R^2_\beta = 0.393$**	
$\mathrm{HR}_{\mathrm{phasic}}$	$\beta = -0.020$ $p = 0.39$ $R^2_\beta = 0.0121$	$\beta = 0.019$ $p = 0.454$ $R^2_\beta = 0.0092$	**$\beta = 0.105$** **$p < 0.001$** **$R^2_\beta = 0.253$**	$\beta = -0.56$ $p = 0.298$ $R^2_\beta = 0.045$

The regression coefficient (β), the p value for the test of $|\beta| > 0$, and the coefficient of determination for fixed effects (R^2_β; Edwards et al., 2008) are displayed. $N = 62$, except for $\mathrm{SCR}_{\mathrm{amp}}$ where $N = 50$, because participants not showing any SCR were excluded from the analysis, and for $\mathrm{SCR}_{\mathrm{lat}}$ where $N = 41$, because $\mathrm{SCR}_{\mathrm{lat}}$ was defined as missing on trials where no SCR occurred. Bold font indicates significant correlations ($p < 0.05$)

criteria defined in the Method section. In the correlation analysis of SCR amplitude, these participants were excluded from the data analysis. Correlation analysis of $\mathrm{SCR}_{\mathrm{lat}}$ was limited to the 401 trials on which—according to the definition provided earlier—a stimulus-related skin conductance response had occurred. In addition, for $\mathrm{SCR}_{\mathrm{lat}}$, participants showing less than three SCRs were removed from the analysis. For this reason, only 41 subjects remained in the analysis of the SCR latency.

As shown in Table 3, the ratings of arousal and valence were negatively correlated. This coincides with the results of Lang et al. (1993) for IAPS pictures and Oberfeld et al. (2009) for colors. As expected, there was also a significant positive correlation between self-rated arousal and the log SCR amplitude. While this is qualitatively consistent with earlier findings (Bradley et al., 2001; Lang et al., 1993), the correlation was weak, according to the classification of Cohen (1988). The previous studies presenting for example emotional pictures reported much higher correlations between self-rated arousal and the log SCR amplitude, for example $r = 0.80$ in Lang et al. (1993). It is important to note, however, that the methods for computing the correlations differed considerably between our study and the study by Lang et al. (1993). In Lang et al. (1993), the arousal ratings for the complete set of 21 emotional IAPS pictures were first rank-ordered within each participant. In the second step, for each of the resulting 21 rank categories, the mean arousal rating and the mean log-transformed SCR amplitude were computed across participants, resulting in 21 pairs of responses that entered the correlation analysis. Thus, Lang et al. analyzed the covariation between the mean affective rating (i.e., valence or arousal) and the mean physiological response at each judgment rank. We conducted the same type of analysis for our set of 30 color stimuli. This resulted in a considerably higher positive correlation between the arousal ratings and log $\mathrm{SCR}_{\mathrm{amp}}$, $r = 0.421$, $p = 0.020$, $n = 30$. As shown in Table 3, the correlation between self-rated valence and log $\mathrm{SCR}_{\mathrm{amp}}$ was not significant (Lang-type correlation: $r = 0.151$, $p = 0.42$, $n = 30$).

Table 3 shows that our data did not exhibit the positive correlation between phasic heart rate and the valence ratings reported by Lang et al. (1993). Even the Lang et al. method for computing the correlation showed no significant association, $r = 0.070$, $p = 0.712$, $n = 30$. We assume that this discrepancy can, in part, be attributed to the restricted range of valence ratings elicited by our color stimuli (see Fig. 2), compared to the wider range observed for the IAPS pictures. The correlation between $\mathrm{HR}_{\mathrm{phasic}}$ and self-rated arousal was also non-significant (see Table 3; Lang-type correlation: $r = -0.103$, $p = 0.588$, $n = 30$).

Within the physiological measures, the SCR latency was significantly negatively correlated with log $\mathrm{SCR}_{\mathrm{amp}}$ (Boucsein, 1992), and the log $\mathrm{SCR}_{\mathrm{amp}}$ was positively correlated with $\mathrm{HR}_{\mathrm{phasic}}$ (see Table 3).

Discussion

To identify effects of color on emotion (valence and arousal), participants viewed a large LED panel (visual angle 18° × 18°) producing light of different colors before a black background. In a factorial design, the color was varied on all of the three perceptual color dimensions (hue, saturation, and brightness), and the resulting 27 chromatic colors were selected using a precise colorimetric specification (CIE L*a*b* color space; Commission Internationale de l'Éclairage, 2007). In addition, three equiluminant gray levels (achromatic colors) were presented. Different to many previous studies, participants were instructed to rate their actual emotions (arousal and valence) induced by the color stimuli, rather than to judge color-emotion associations. In addition, skin conductance and heart rate were measured during color presentation.

All the three color dimensions (hue, saturation, and brightness) had a significant effect on the emotion ratings. Across the 30 colors presented in the experiment, changes in color had a stronger effect on arousal than on valence (see Fig. 2). The variation in arousal was only slightly smaller than the range of arousal ratings obtained for the International Affective Picture System (IAPS; Lang et al., 1998).

As expected, the color dimension saturation had a strong effect on self-rated arousal, with higher saturation corresponding to higher arousal ratings, which is consistent with results by Valdez and Mehrabian (1994). The pattern is also compatible with data by Suk and Irtel (2010), but it should be noted that they analyzed the arousal ratings as a function of chroma which, unlike saturation, is related to both "color purity" and brightness (Fairchild, 2005). Thus, the positive correlation between arousal and chroma reported by Suk and Irtel (2010) can likely be attributed to effects of both saturation and brightness on arousal. In a recent study, arousal ratings were obtained for color patches varying in chromaticness (saturation), brightness, and hue (blue, green, yellow, and red) according to the Natural Color System (NCS) (Zielinski, 2016). Unfortunately, the paper does not specify whether the stimuli were presented on a color-calibrated display, and colorimetric values are not reported. For these reasons, it is difficult to judge whether Zielinski (2016) did indeed present a factorial combination of hue, saturation, and brightness as in the present study. Still, the report of a significant positive relation between saturation and self-rated arousal is compatible with our data (Zielinski, 2016).

Our data show that when controlling for saturation and brightness, red was the most arousing color, followed by green and blue. This is in accordance with studies by Suk and Irtel (2010), Oberfeld et al. (2009), Yildirim, Hidayetoglu, and Capanoglu (2011) and Zielinski (2016), and also with older studies that often did not consider potential effects of saturation and brightness (e.g., Gerard, 1958; Wilson, 1966). A deviating result was reported by Valdez and Mehrabian (1994) who observed highest arousal ratings for a green–yellow hue. The fact that gray received the lowest average arousal ratings can be attributed to an effect of saturation.

It is difficult to judge how our observation that higher brightness caused higher arousal ratings aligns with results from the previous studies. Valdez and Mehrabian (1994) reported a U-shaped relationship between brightness and arousal. They presented chromatic colors with brightness values between 10 and 60 in terms of the Munsell color system (small color patches viewed under light with a correlated color temperature of 5000 K), and achromatic colors with brightness values between 5 and 80. The arousal ratings decreased with brightness up to a brightness value of about 45 and then increased again. Unfortunately, Valdez and Mehrabian (1994) did not report colorimetric measurements (nor the exact Munsell color values) for their stimuli, so that it is not possible to compare the luminance range of their stimuli to our study. In addition, it is unlikely that they varied brightness and saturation independently as in the present study, so that it remains unclear to which extent the effect of brightness on arousal is confounded by differences in saturation. Suk and Irtel (2010) presented achromatic stimuli with lightness levels varying from $L^* = 0$ (black) to 100 (white) and also observed a u-shaped relation between lightness and arousal. As they did not report luminance measurements, it is again not possible to compare the luminance range of our stimuli to their study. The same problem applies to data by Zielinski (2016) who presented chromatic colors and reported no significant effect of brightness on arousal.

Apart from these main effects of the three color dimensions on the arousal ratings, the data show several significant interactions. Most important, the hue and the brightness affected the arousal only for colors with relatively high saturation (see Fig. 3a, c). These interactions highlight the importance of presenting a well-defined set of color stimuli varying across all of the three perceptual color dimensions.

For the valence ratings, the largest difference was observed between chromatic and achromatic colors. On average, presentation of chromatic colors led to more positive ratings than the presentation of achromatic stimuli, consistent with results by Suk and Irtel (2010). This difference in valence was larger in female than in male participants. For the chromatic stimuli, compatible with data by Valdez and Mehrabian (1994), Suk and Irtel (2010) and Oberfeld et al. (2009), there was a trend towards an effect of hue on the valence ratings. The data showed the increase in valence from green over red to blue reported previously, but only at the highest saturation level. At low saturation, this pattern was reversed, again demonstrating that not the hue alone but rather the combination of hue and the other two perceptual dimensions determines the emotional effects of a color. As in Valdez and Mehrabian (1994) and

Suk and Irtel (2010), our data showed a significant main effect of saturation on the valence ratings. However, in the two previous studies, valence showed a monotonic increase with saturation, while in our data, the highest average valence ratings (across the three chromatic hues) were obtained at medium rather than at high saturation. In addition, the effect of saturation differed between the three hues (see Fig. 4a). For green and blue, the valence ratings at high and medium saturation were higher than at low saturation. For red, the highest valence rating was observed at medium saturation, followed by low and then high saturation. As expected on the basis of Valdez and Mehrabian's (1994) findings, we observed an increase of the valence ratings with increasing brightness. For achromatic colors, Suk and Irtel (2010) reported a u-shaped pattern, with the highest valance ratings at low and high brightness levels for achromatic stimuli. As mentioned above, it is difficult to compare their results to our study, due to the unknown luminance of the stimuli in Suk and Irtel (2010). Taken together, as for arousal, the self-rated valence of participants' emotional state while viewing colors was affected by hue, saturation, brightness, and their interactions.

It is interesting to compare the observed effects of color on self-rated valence with studies on color preferences. A preferred color could be viewed as having a higher valence than a non-preferred color, although it is of course an important difference that in color preference studies, the *colors* are judged, while in our study, participants rated their own emotional state while viewing a certain color. The rank order of valence ratings observed for saturated colors (see Fig. 4a) corresponds to the results of many studies on color preference in western cultures, which show the highest rank for blue, followed by green or red (e.g., Crozier, 1999; Eysenck, 1941; Guilford & Smith, 1959). Effects of saturation and brightness on color preference have also been reported (e.g., Camgöz, Yener, & Guvenc, 2002; Guilford & Smith, 1959; Hogg, Goodman, Porter, Mikellides, & Preddy, 1979; Ou, Luo, Woodcock, & Wright, 2004b; Tate & Allen, 1985). In our data, the variation in self-rated valence appears somewhat weaker than the variation in 'color liking' found in some color preference studies. We also observed several interaction effects between hue, brightness, and saturation. In some studies on color preference, the rank order of the hues blue, green, and red depended on brightness and saturation (Guilford & Smith, 1959; Palmer & Schloss, 2010), while it was preserved in other studies (e.g., McManus, Jones, & Cottrell, 1981). Apart from the basic difference between judgments of color liking and ratings of the participant's own emotional state in terms of valence, the differences between our data and results from color preference studies could be due to stimulus-related factors, such as the relatively large viewing angle or relatively long viewing time used in our experiment.

Did the effects of color that we observed for the experiential component of emotion also show for physiological emotional responses (cf. Mauss & Robinson, 2009)? Not surprisingly, the effects of color on skin conductance and heart rate were weaker than the effects on the emotion ratings (Lench, Flores, & Bench, 2011). However, for chromatic colors, saturation and brightness had a significant effect on the SCR amplitude, which is supposed to be a physiological measure of arousal (Lang et al., 1993), even though we did not exclude the 12 (out of 62) participants who did not produce any SCRs across the 30 colors. The color dimension saturation had the strongest influence on the SCR amplitude, compatible with results by Zielinski (2016). Higher saturation resulted in larger SCR amplitudes, similar to the arousal ratings (Fig. 3a). However, comparatively high SCR amplitudes were also observed for the gray levels (zero saturation), which contradicts the monotonic positive relation between saturation and SCR amplitude suggested by the data for the chromatic colors and does not match the arousal ratings. The observed (nonsignificant) increase of the SCR amplitude from blue to green to red is compatible with the arousal ratings (Fig. 3a), and also with the previous studies (Gerard, 1958; Jacobs & Hustmyer, 1974; Wilson, 1966) that did not control for brightness and saturation. Zielinski (2016) reported no significant effect of hue on the SCR. Our data also show a significant positive relation between brightness and $\mathrm{SCR}_{\mathrm{amp}}$, again compatible with the arousal ratings (Fig. 3b). This effect has not been investigated systematically in most previous studies measuring electrodermal activity. Zielinski (2016) reported no significant effect of brightness on the SCR, which might be due to a different range of presented brightness levels. However, as noted above, we cannot directly compare the colors between his and our study, because Zielinski (2016) did not specify colorimetric values. The latency of the SCR responses showed no significant effects of color.

The phasic, stimulus-related change in heart rate (Fig. 6) showed a similar pattern as the valence ratings (Fig. 4a). This is the expected result given the positive correlation between an acceleration in the HR and valence ratings for IAPS pictures reported by Lang et al. (1993). However, the only significant effect of color on $\mathrm{HR}_{\mathrm{phasic}}$ was due to the achromatic stimuli, which, on average, caused a deceleration of the heart rate, while the chromatic stimuli resulted in an acceleration. Still, for chromatic colors, there was a trend towards a stronger HR acceleration at higher brightness levels, compatible with the pronounced effect of brightness on the valence ratings (Fig. 4b).

Compatible with data obtained for the IAPS picture set (Lang et al., 1993), our results show a significant positive correlation between self-rated arousal and the SCR amplitude. However, the correlation was weak. Even when using the specific method by Lang et al. (1993) for computing the correlation, which produces much higher correlation coefficients than the within-subjects correlation approach by Edwards et al. (2008), we obtained a smaller correlation coefficient ($r = 0.42$) than Lang et al. ($r = 0.80$). Zielinski (2016) also found a slightly higher correlation ($r = 0.64$) between self-rated arousal and SCR amplitude with the Lang et al. method, but he presented only 15 different colors rather than 30 colors as in the present study. The positive correlation between valence ratings and stimulus-induced changes in heart rate reported by Lang et al. (1993) was not present in our data. One reason for these differences could be that color presentation represents a weaker emotion-relevant stimulation than presenting IAPS pictures, which according to a meta-analysis by Lench et al. (2011) are relatively effective in inducing emotions. Still, the emotion ratings and physiological responses obtained in our study demonstrate that color can be added to the list of emotion-inducing stimuli. In fact, using color to elicit different levels of arousal could be advantageous in emotion research when it is desirable to conceal the intent to modify the emotional state of the observer.

Our results clearly demonstrate that the effect of a given color on emotion is not determined by, for example, the hue alone, but by the combination of hue, saturation, and brightness. For this reason, when confronted with statements such as “Red causes higher arousal than blue”, one should ask “Which red and which blue?”. Figure 3 shows that while the statement concerning the effects of red versus blue on arousal is correct when saturation and brightness are controlled for, it is easy to select a color with a blue hue (e.g., highly saturated and bright) that corresponds to higher arousal than a color with a red hue (e.g., lowly saturated and dim). Our data have implications for the use of color in applied design contexts, and when using colors to communicate a certain meaning or to induce a certain mood: it is important to consider the interactions between hue, saturation, and brightness.

Which could be the mechanisms underlying the effects of hue, saturation, and brightness on emotion? Concerning the relatively strong effect of color on arousal evidenced by both the emotion ratings and the SCR amplitudes in our study, one potential explanation involves the melatonin-suppressing retinal ganglion cells (e.g., Berson, Dunn, & Takao, 2002; Brainard et al., 2001; Provencio et al., 2000). These photosensitive cells in the human retina project to the suprachiasmatic nuclei which are involved the regulation of circadian transitions between sleep and wakefulness (Aston-Jones, 2005; Moore, 1983). The latter two states can be viewed as different levels of arousal. In fact, the peak in the wavelength spectrum of the blue primary of our LED display is close to the peak sensitivity in the action spectrum of the Melanopsin-expressing ganglion cells (Brainard et al., 2001; Thapan, Arendt, & Skene, 2001). However, activation of these photoreceptors by short-wavelength light should result in an increase in alertness or arousal (e.g., Yasukouchi & Ishibashi, 2005), while our data show significantly lower arousal under short-wavelength blue light compared to longer-wavelength red light. In principle, this seemingly contradictory effect could be due to the modulating input of S-, M-, and L-cones (Kaiser & Boynton, 1996) to the Melanopsin-expressing ganglion cells (Dacey et al., 2005), where S-cone input is inhibitory, while L- and M-cone input is excitatory. However, a quantitative model of the activity of Melanopsin-expressing ganglion cells would be necessary to answer the question if these cone inputs could explain the effects of hue on arousal observed in our study. In addition, our results show a stronger effect of the saturation than of the hue of a color on arousal, and this pattern is difficult to explain with the activation of a single photoreceptor system. Thus, it remains for future research to identify the mechanisms responsible for emotional effects of color. On a more general level, it remains to be shown if the effects of color on emotion are predominantly caused by cognitive appraisal, for instance due to color-meaning associations (Adams & Osgood, 1973), or arise because of more direct (or less cognitive) mechanisms.

Although compared to most previous studies, our experiment introduced important methodological advances, most importantly, the presentation of colorimetrically specified colored light with independent variation of hue, saturation, and brightness in a factorial design, there are of course limitations. It would be desirable to measure emotional effects across a wider range of colors in future experiments (e.g., other hues such as yellow or purple, and wider ranges of brightness and saturation). In addition, we presented a comparably large self-illuminated color stimulus (18° × 18° of visual angle) before a black background, and it remains to be shown if our results generalize to other contexts like small passively illuminated color patches before achromatic or chromatic backgrounds, colored interior spaces, or colored clothing.

Concerning measures of the subjective experience of emotion, we used a dimensional approach and obtained ratings of arousal and valence. It would be desirable to study the effects of our set of colors on the emotion dimension dominance/potency (e.g., Suk & Irtel, 2010), or on discrete emotions like joy, anger, or shame (e.g., Scherer, 2005). In a recent study, Dael, Perseguers, Marchand, Antonietti, and Mohr (2016) showed videos of professional actors expressing either panic fear or elated joy. Participants were asked to select a color matching the expressed emotion. Across subjects, the best-matching colors for the two emotions differed with respect to all of the three perceptual dimensions of color (hue, saturation, and brightness), compatible with the results of the present study. Also compatible with the valence ratings obtained in our study (Fig. 4), the colors matched to joy expressions were brighter and more saturated than the colors matched to fear expressions. It remains to be shown, however, if these differences are observed not only for discrete emotions differing strongly in valence as in the study by Dael et al., but also for example for two emotions from the positive spectrum (e.g., pride versus admiration).

To summarize, our study confirms that color has systematic effects on the emotional state of a person viewing the color. The effect on arousal and valence as well as on skin conductance responses and heart rate is determined by the specific combination of hue, saturation, and brightness constituting a color.

Acknowledgements We are grateful to Gerhard Vossel for helpful suggestions concerning the experimental setup and the registration of the physiological parameters. We thank Agnes Münch for her valuable help with the technical implementation of the experiment and with the analysis of the physiological data. We are grateful to Martin Happ for extending his HRM R package (https://cran.r-project.org/package=HRM), so that we could use it for analyses involving three within-subjects factors. Two anonymous reviewers provided helpful comments on an earlier version of this paper.

Compliance with ethical standards

Funding No external funding received.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Adams, F. M., & Osgood, C. E. (1973). A cross-cultural study of the affective meanings of color. *Journal of Cross-Cultural Psychology, 4*(2), 135–156.

Ali, M. R. (1972). Pattern of EEG recovery under photic stimulation by light of different colors. *Electroencephalography and Clinical Neurophysiology, 33*(3), 332–335.

Arnau, J., Bono, R., & Vallejo, G. (2009). Analyzing small samples of repeated measures data with the mixed-model adjusted F Test. *Communications in Statistics-Simulation and Computation, 38*(5), 1083–1103. doi:10.1080/03610910902785746.

Ashcroft, K. R., Guimaraes, F. S., Wang, M., & Deakin, J. F. W. (1991). Evaluation of a psychophysiological model of classical fear conditioning in anxious patients. *Psychopharmacology (Berl), 104*(2), 215–219. doi:10.1007/Bf02244181.

Aston-Jones, G. (2005). Brain structures and receptors involved in alertness. *Sleep Medicine, 6,* S3–S7. doi:10.1016/s1389-9457(05)80002-4.

Berson, D. M., Dunn, F. A., & Takao, M. (2002). Phototransduction by retinal ganglion cells that set the circadian clock. *Science, 295*(5557), 1070–1073. doi:10.1126/science.1067262.

Bland, J. M., & Altman, D. G. (1995a). Statistics notes: 12. Calculating correlation coefficients with repeated observations: Part 1: Correlation within subjects. *British Medical Journal, 310*(6977), 446.

Bland, J. M., & Altman, D. G. (1995b). Statistics notes: 13. Calculating correlation coefficients with repeated observations: Part 2: Correlation between subjects. *British Medical Journal, 310*(6980), 633.

Boucsein, W. (1992). *Electrodermal activity*. New York: Plenum Press.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology, 49*(8), 1017–1034. doi:10.1111/j.1469-8986.2012.01384.x.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion, 1*(3), 276–298. doi:10.1037//1528-3542.1.3.276.

Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). Remembering pictures: Pleasure and arousal in memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 18*(2), 379–390.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment mannequin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49–59. doi:10.1016/0005-7916(94)90063-9.

Brainard, D. H. (2003). Color appearance and color difference specification. In S. K. Shevell (Ed.), *The science of color* (2nd ed., pp. 191–216). Amsterdam, Boston: Elsevier & Optical Society of America.

Brainard, G. C., Hanifin, J. P., Greeson, J. M., Byrne, B., Glickman, G., Gerner, E., & Rollag, M. D. (2001). Action spectrum for melatonin regulation in humans: Evidence for a novel circadian photoreceptor. *Journal of Neuroscience, 21*(16), 6405–6412.

Burton, P., Gurrin, L., & Sly, P. (1998). Extending the simple linear regression model to account for correlated responses: An introduction to generalized estimating equations and multi-level mixed modelling. *Statistics in Medicine, 17*(11), 1261–1291.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 173–191). New York: The Guilford Press.

Camgöz, N., Yener, C., & Guvenc, D. (2002). Effects of hue, saturation, and brightness on preference. *Color Research and Application, 27*(3), 199–207. doi:10.1002/Col.10051.

Carterette, E. C., & Symmes, D. (1952). Color as an experimental variable in photic stimulation. *Electroencephalography and Clinical Neurophysiology, 4*(3), 289–296. doi:10.1016/0013-4694(52)90054-0.

Carvalho, S., Leite, J., Galdo-Alvarez, S., & Goncalves, O. F. (2012). The emotional movie database (EMDB): A self-report and psychophysiological study. *Applied Psychophysiology and Biofeedback, 37*(4), 279–294. doi:10.1007/s10484-012-9201-6.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J.: L. Erlbaum Associates.

Commission Internationale de l'Éclairage (2006). Colorimetry. Part 1: CIE standard colorimetric observers (Vol. CIE S 014-1/E:2006/ ISO 11664-1:2008(E)).

Commission Internationale de l'Éclairage (2007). Colorimetry. Part 4: CIE 1976 L*a*b* colour spaces (Vol. CIE S 014-4/E:2007 (ISO 11664-4:2008)).

Commission Internationale de l'Éclairage (2008). Colorimetry: Part 2: Standard llluminants for colorimetry (Vol. CIE S 014-2/ E:2006 (ISO 11664-2:2007(E))).

Crozier, W. R. (1999). The meanings of colour: preferences among hues. *Pigment & Resin Technology, 28*(1), 6–14. doi:10.1108/03699429910252315.

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biological Psychology, 52*(2), 95–111.

Dacey, D. M., Liao, H. W., Peterson, B. B., Robinson, F. R., Smith, V. C., Pokorny, J., & Gamlin, P. D. (2005). Melanopsin-expressing ganglion cells in primate retina signal colour and irradiance and project to the LGN. *Nature, 433*(7027), 749–754. doi:10.1038/nature03387.

Dael, N., Perseguers, M. N., Marchand, C., Antonietti, J. P., & Mohr, C. (2016). Put on that colour, it fits your emotion: Colour appropriateness as a function of expressed emotion. *Quarterly Journal of Experimental Psychology, 69*(8), 1619–1630. doi:10.1080/17470218.2015.1090462.

Edwards, L. J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger, O. (2008). An R^2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine, 27*(29), 6137–6157. doi:10.1002/sim.3429.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*(3–4), 169–200. doi:10.1080/02699939208411068.

Elliot, A. J., & Maier, M. A. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology, 65,* 95–120. doi:10.1146/annurev-psych-010213-115035.

Eysenck, H. J. (1941). A critical and experimental study of colour preferences. *The American Journal of Psychology, 54*(3), 385–394.

Fairchild, M. D. (2005). *Color appearance models*. Chichester, UK: Wiley.

Gao, X. P., Xin, J. H., Sato, T., Hansuebsai, A., Scalzo, M., Kajiwara, K., & Billger, M. (2007). Analysis of cross-cultural color emotion. *Color Research and Application, 32*(3), 223–229. doi:10.1002/col.20321.

Gerard, R. M. (1958). Color and emotional arousal. *American Psychologist, 13*(7), 340.

Gomez, P., Zimmermann, P., Guttormsen-Schar, S., & Danuser, B. (2005). Respiratory responses associated with affective processing of film stimuli. *Biological Psychology, 68*(3), 223–235. doi:10.1016/j.biopsycho.2004.06.003.

Greene, T. C., Bell, P. A., & Boyer, W. N. (1983). Coloring the environment: Hue, arousal, and boredom. *Bulletin of the Psychonomic Society, 21*(4), 253–254.

Guilford, J. P., & Smith, P. C. (1959). A system of color preferences. *American Journal of Psychology, 72*(4), 487–502. doi:10.2307/1419491.

Happ, M., Harrar, S. W., & Bathke, A. C. (2016). Inference for low- and high-dimensional multigroup repeated measures designs with unequal covariance matrices. *Biometrical Journal, 58*(4), 810–830. doi:10.1002/bimj.201500064.

Happ, M., Harrar, S. W., & Bathke, A. C. (2017). High-dimensional repeated measures. *Journal of Statistical Theory and Practice*. doi:10.1080/15598608.2017.1307792.

Heller, E. (2008). *Wie Farben wirken: Farbpsychologie, Farbsymbolik, kreative Farbgestaltung* (4th ed.). Reinbek: Rowohlt.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika, 75*(4), 800–802.

Hogg, J., Goodman, S., Porter, T., Mikellides, B., & Preddy, D. E. (1979). Dimensions and determinants of judgements of colour samples and a simulated interior space by architects and non-architects. *British Journal of Psychology, 70*(2), 231–242.

Hurlbert, A. C., & Ling, Y. Z. (2007). Biological components of sex differences in color preference. *Current Biology, 17*(16), R623–R625. doi:10.1016/j.cub.2007.06.022.

Ishihara, S. (2013). *The series of plates designed as a test of colour deficiency. 24* (Plates ed.). Tokyo: Kanehara Trading Inc.

Jacobs, K. W., & Hustmyer, F. E. (1974). Effects of four psychological primary colors on GSR, heart-rate and respiration rate. *Perceptual and Motor Skills, 38*(3), 763–766.

Jameson, D., Hurvich, L. M., & Varner, F. D. (1979). Receptoral and post-receptoral visual processes in recovery from chromatic adaptation. *Proceedings of the National Academy of Sciences of the United States of America, 76*(6), 3034–3038.

Kaiser, P. K. (1984). Physiological response to color: A critical review. *Color Research and Application, 9*(1), 29–36.

Kaiser, P. K., & Boynton, R. M. (1996). *Human color vision* (2nd ed.). Washington, DC: Optical Society of America.

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics, 53*(3), 983–997. doi:10.2307/2533558.

Keselman, H. J., Algina, J., & Kowalchuk, R. K. (2001). The analysis of repeated measures designs: A review. *British Journal of Mathematical and Statistical Psychology, 54,* 1–20.

Keselman, H. J., Algina, J., & Kowalchuk, R. K. (2002). A comparison of data analysis strategies for testing omnibus effects in higher-order repeated measures designs. *Multivariate Behavioral Research, 37*(3), 331–357. doi:10.1207/S15327906mbr3703_2.

Knutson, B., Katovich, K., & Suri, G. (2014). Inferring affect from fMRI data. *Trends in Cognitive Sciences, 18*(8), 422–428. doi:10.1016/j.tics.2014.04.006.

Kowalchuk, R. K., Keselman, H. J., Algina, J., & Wolfinger, R. D. (2004). The analysis of repeated measurements with mixed-model adjusted *F* tests. *Educational and Psychological Measurement, 64*(2), 224–242. doi:10.1177/0013164403260196.

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology, 84*(3), 394–421. doi:10.1016/j.biopsycho.2010.03.010.

Küller, R., Mikellides, B., & Janssens, J. (2009). Color, arousal, and performance: A comparison of three experiments. *Color Research and Application, 34*(2), 141–152. doi:10.1002/col.20476.

Lane, R. D., Reiman, E. M., Bradley, M. M., Lang, P. J., Ahern, G. L., Davidson, R. J., & Schwartz, G. E. (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia, 35*(11), 1437–1444. doi:10.1016/s0028-3932(97)00070-5.

Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in Mental Health Care Delivery Systems* (pp. 119–137). Norwood, NJ: Ablex.

Lang, P. J., Bradley, M. M., Fitzsimmons, J. R., Cuthbert, B. N., Scott, J. D., Moulder, B., & Nangia, V. (1998a). Emotional arousal and activation of the visual cortex: An fMRI analysis. *Psychophysiology, 35*(2), 199–210.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261–273. doi:10.1111/j.1469-8986.1993.tb03352.x.

Lang, P. J., Öhman, A., & Vaitl, D. (1998b). *The international affective picture system*. Gainesville, FL: Univeristy of Florida, Center for Research. **(in Psychophysiology)**.

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin, 137*(5), 834–855. doi:10.1037/a0024244.

Luce, R. D. (1991). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.

Mauss, I., & Robinson, M. (2009). Measures of emotion: A review. *Cognition and Emotion, 23*(2), 209–237. doi:10.1080/02699930802204677.

McManus, I. C., Jones, A. L., & Cottrell, J. (1981). The aesthetics of colour. *Perception, 10*(6), 651–666. doi:10.1068/p100651.

Meerwein, G., Rodeck, B., & Mahnke, F. H. (2007). *Farbe: Kommunikation im Raum* (4th ed.). Basel: Birkhäuser.

Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. Cambridge: M.I.T. Press.

Mikellides, B. (1990). Color and physiological arousal. *Journal of Architectural and Planning Research, 7*(1), 13–20.

Moore, R. Y. (1983). Organization and function of a central nervous-system circadian oscillator: The suprachiasmatic hypothalamic nucleus. *Federation Proceedings, 42*(11), 2783–2789.

Nourse, J. C., & Welch, R. B. (1971). Emotional attributes of color: Comparison of violet and green. *Perceptual and Motor Skills, 32*(2), 403–406.

Oberfeld, D., Hecht, H., Allendorf, U., & Wickelmaier, F. (2009). Ambient lighting modifies the flavor of wine. *Journal of Sensory Studies, 24*(6), 797–832. doi:10.1111/j.1745-459X.2009.00239.x.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Oxford, England: Univer. Illinois Press.

Ou, L. C., Luo, M. R., Woodcock, A., & Wright, A. (2004a). A study of colour emotion and colour preference. Part I: Colour emotions for single colours. *Color Research and Application, 29*(3), 232–240. doi:10.1002/Col.20010.

Ou, L. C., Luo, M. R., Woodcock, A., & Wright, A. (2004b). A study of colour emotion and colour preference. Part III: Colour preference modeling. *Color Research and Application, 29*(5), 381–389.

Palmer, S. E., & Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proceedings of the National academy of Sciences of the United States of America, 107*(19), 8877–8882. doi:10.1073/pnas.0906172107.

Palomba, D., Sarlo, M., Angrilli, A., Mini, A., & Stegagno, L. (2000). Cardiac responses associated with affective processing of unpleasant film stimuli. *International Journal of Psychophysiology, 36*(1), 45–57. doi:10.1016/S0167-8760(99)00099-9.

Provencio, I., Rodriguez, I. R., Jiang, G. S., Hayes, W. P., Moreira, E. F., & Rollag, M. D. (2000). A novel human opsin in the inner retina. *Journal of Neuroscience, 20*(2), 600–605.

R Core Team (2017). R: A language and environment for statistical computing. Wien: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Robertson, A. R. (1990). Historical development of CIE recommended color difference equations. *Color Research and Application, 15*(3), 167–170. doi:10.1002/col.5080150308.

Roedema, T. M., & Simons, R. F. (1999). Emotion-processing deficit in alexithymia. *Psychophysiology, 36*(3), 379–387. doi:10.1017/S0048577299980290.

Rouanet, H., & Lépine, D. (1970). Comparison between treatments in a repeated-measurement design: ANOVA and multivariate methods. *British Journal of Mathematical and Statistical Psychology, 23*(2), 147–163.

Sato, T., Kajiwara, K., Hoshino, H., & Nakamura, T. (2000). Quantitative evaluation and categorising of human emotion induced by colour. *Advances in Colour Science and Technology, 3*, 53–59.

Schaie, K. W. (1961). Scaling the association between colors and mood-tones. *American Journal of Psychology, 74*(2), 266. doi:10.2307/1419412.

Schauss, A. G. (1985). The physiological effect of color on the suppression of human aggression: Research on Baker-Miller pink. *International Journal of Biosocial Research, 7*(2), 55–64.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information Sur Les Sciences Sociales, 44*(4), 695–729. doi:10.1177/0539018405058216.

Soriano, C., & Valenzuela, J. (2009). Emotion and colour across languages: implicit associations in Spanish colour terms. *Social Science Information Sur Les Sciences Sociales, 48*(3), 421–445. doi:10.1177/0539018409106199.

Storm, H., Fremming, A., Odegaard, S., Martinsen, O. G., & Morkrid, L. (2000). The development of a software program for analyzing spontaneous and externally elicited skin conductance changes in infants and adults. *Clinical Neurophysiology, 111*(10), 1889–1898. doi:10.1016/s1388-2457(00)00421-1.

Suk, H. J., & Irtel, H. (2010). Emotional response to color across media. *Color Research and Application, 35*(1), 64–77. doi:10.1002/col.20554.

Tate, F. B., & Allen, H. (1985). Color preferences and the aged individual: Implications for art therapy. *Arts in Psychotherapy, 12*(3), 165–169. doi:10.1016/0197-4556(85)90016-4.

Terwogt, M. M., & Hoeksma, J. B. (1995). Colors and emotions: Preferences and combinations. *Journal of General Psychology, 122*(1), 5–17.

Thapan, K., Arendt, J., & Skene, D. J. (2001). An action spectrum for melatonin suppression: evidence for a novel non-rod, non-cone photoreceptor system in humans. *Journal of Physiology-London, 535*(1), 261–267. doi:10.1111/j.1469-7793.2001.t01-1-00261.x.

Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology: General, 123*(4), 394–409.

Vandewalle, G., Schwartz, S., Grandjean, D., Wuillaume, C., Balteau, E., Degueldre, C., & Maquet, P. (2010). Spectral quality of light modulates emotional brain responses in humans. *Proceedings of the National academy of Sciences of the United States of America, 107*(45), 19549–19554. doi:10.1073/pnas.1010180107.

Vary, P., & Stiny, L. (1977). Ein Herz-Rhythmus-Analysator für Langzeit-EKG-Untersuchungen [A Heart Rate Analyzer for Long-Term ECG Recordings]. *Biomedizinische Technik, 22*(3), 39–44. doi:10.1515/bmte.1977.22.3.39.

Walters, J., Apter, M. J., & Svebak, S. (1982). Color preference, arousal, and the theory of psychological reversals. *Motivation and Emotion, 6*(3), 193–215.

Whitfield, T. W. A., & Wiltshire, T. J. (1990). Color psychology: A critical review. *Genetic, Social, and General Psychology Monographs, 116*(4), 385–411.

Wilson, G. D. (1966). Arousal properties of red versus green. *Perceptual and Motor Skills, 23*(3, Part 1), 947–949.

Witvliet, C. V., & Vrana, S. R. (1995). Psychophysiological responses as indexes of affective dimensions. *Psychophysiology, 32*(5), 436–443.

Wolfinger, R. D. (1996). Heterogeneous variance: Covariance structures for repeated measures. *Journal of Agricultural, Biological, and Environmental Statistics, 1*(2), 205–230.

Wright, B., & Rainwater, L. (1962). The meanings of color. *Journal of General Psychology, 67*, 89–99.

Wundt, W. M. (1903). *Grundriss der physiologischen Psychologie*. Leipzig: Engelmann.

Wyszecki, G., & Stiles, W. S. (2000). *Color science: Concepts and methods, quantitative data, and formulae (Wiley Classics Library edn.)*. New York: Wiley.

Yasukouchi, A., & Ishibashi, K. (2005). Non-visual effects of the color temperature of fluorescent lamps on physiological aspects in humans. *J Physiol Anthropol Appl Human Sci, 24*(1), 41–43. **(JST.JSTAGE/jpa/24.41)**.

Yildirim, K., Hidayetoglu, M. L., & Capanoglu, A. (2011). Effects of interior colors on mood and preference: comparisons of two living rooms. *Perceptual and Motor Skills, 112*(2), 509–524. doi:10.2466/24.27.pms.112.2.509-524.

Zielinski, P. (2016). An arousal effect of colors saturation: A study of self-reported ratings and electrodermal responses. *Journal of Psychophysiology, 30*(1), 9–16. doi:10.1027/0269-8803/a000149.