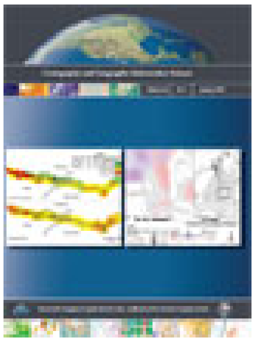


Taylor & Francis
Taylor & Francis Group

Cartography and Geographic Information Science

ISSN: 1523-0406 (Print) 1545-0465 (Online) Journal homepage: http://www.tandfonline.com/loi/tcag20

The effect of spatial distance on the discriminability of colors in maps

Alžběta Brychtová & Arzu Çöltekin

To cite this article: Alžběta Brychtová & Arzu Çöltekin (2016): The effect of spatial distance on the discriminability of colors in maps, Cartography and Geographic Information Science, DOI: 10.1080/15230406.2016.1140074

To link to this article: http://dx.doi.org/10.1080/15230406.2016.1140074

Published online: 15 Feb 2016.

http://dx.doi.org/10.1080/15230406.2016.1140074

Taylor & Francis
Taylor & Francis Group

The effect of spatial distance on the discriminability of colors in maps

Alžběta Brychtová[a] and Arzu Çöltekin[b]

[a]Department of Geoinformatics, Faculty of Science, Palacký University, Olomouc, Czech Republic; [b]Department of Geography, Faculty of Science, University of Zurich, Zurich, Switzerland

ABSTRACT

The spatial distance (gap) between map symbols can have a great impact on their discriminability, however, there is little empirical evidence to establish spatial and attribute thresholds. In this paper, we examine the effect of the spatial gap in discriminability of color *hue* and *value*, that is, we conducted an online study to obtain performance metrics; then an eye-tracking study to understand participants' strategies and cognitive processes. Participants completed two experimental tasks (compare two areas and decide if their color is the same; and compare three areas and rank them from the lightest to the darkest). The *color distances* and the *spatial distances* were strictly controlled for the compared areas. Our analyses confirmed that, overall, increasing the gap between colors has a consistent negative impact on the ability to differentiate them with both sequential and qualitative schemes. Furthermore, we observed that sequential schemes require larger color distances than qualitative schemes for discriminability. Finally, our results suggested that for qualitative colors, the largest tested color distance ΔE_{00} = 10 yields considerably higher levels of accuracy in color discrimination (even when the spatial gap between the two colors is large), thus we recommend ΔE_{00} = 10 to practicing cartographers and other information visualization designers.

ARTICLE HISTORY
Received 7 July 2015
Accepted 6 January 2016

KEYWORDS
cartography; color schemes; spatial distance; color distance; usability; eye-tracking

Introduction and background

In cartography, color is used to depict important information. Therefore, being able to match or discriminate colors is very important for visually identifying information, such as spatial patterns or anomalies. This visual identification of patterns and anomalies depends on whether one can perceive if the two shades or colors are the same or not. Successful matching or discriminating colors depends on design choices (the chromatic differences and the spatial gap between colors), as well as individual differences in perceptual abilities.

Human color perception is highly dependent on environmental as well as psychophysical (biological and cognitive) factors (May 2009). There are strong differences in the way humans experience color (Asano et al. 2015). For example, the perception of color is strongly affected by various factors such as the amount of the light in the environment, objects casting shadows, surrounding materials and their reflectivity as well as observers' previous knowledge and cognitive biases (Derefeldt et al. 2004; Foster 2011). Furthermore, it is well documented that the number and distribution of photoreceptors in the eye influences what we see (e.g., aging and color deficiencies, see Roy et al. 1991) as well as (arguably) our brain assuming certain light direction or source (e.g., Gegenfurtner, Bloj, and Toscani 2015; Lafer-Sousa, Hermann, and Conway 2015; Winkler et al. 2015). In summary, we understand that color perception is not *stable* over space and time for one individual; nor is it between individuals or groups.

In cartography, the importance of color has been long acknowledged both from designers' and map readers' perspectives (Brewer et al. 1997; Brewer 1994; Olson 1987). Color is frequently used to represent categories and other important information, and thus rely on human perception to detect patterns for visual analysis, or to identify the difference between categories for precise comparison tasks (e.g., Dall'Acqua, Çöltekin, and Noetzli 2013). Successfully executing such tasks is essential for map-based decision-making; and while some of these decisions might be viewed as trivial (everyday navigation), some can be life threatening (e.g., sea navigation, emergency, and rescue operations) or costly (e.g., climate change mitigation decisions) (Sheppard et al. 2008; Vande Velde et al. 2009).

CONTACT Alžběta Brychtová alzbeta.brychtova@gmail.com

For successful transmission of spatial information, map users need to be able to identify the meaning of the map symbols, compare them with a legend, and distinguish them from other symbols (Bjørke 1996). Cartographers address these needs typically through making informed design choices; for example, visual hierarchy and visual complexity is controlled through generalization procedures, and appropriate visual variables are modified to suppress or emphasize information (Bertin 1983). Color is such a visual variable, and it is largely utilized in cartography, geovisualization and other forms of information visualization. Indeed, the originally proposed seven *visual* (or retinal) variables by Bertin (1983) include two color-related variables: *position, size, shape*, ***color hue, color value***, *orientation,* and *texture*.

Various researchers examined the effect of visual variables on human performance with maps. The saliency (visual dominance and noticeability) of map symbols based on their *size* was studied, for example, by Meihoefer (1969), Gilmartin (1981) or Dent, Torguson, and Hodler (2009). Furthermore, *size* as well as *orientation* was examined by Garlandini and Fabrikant (2009) and *orientation* alone by Cybulski (2014). Even though there are not many empirical studies evaluating the impact of visual variables, *color hue and value* are arguably the most frequently studied ones. One of the central questions about color in cartography has been on finding the optimal color schemes for choropleth maps (Al-Ghamdi 2014; Cromley 1989; Mersey 1990). Furthermore, cartographers have been interested in studying color deficiencies and/or color blindness; and have developed widely used tools to help designers to choose colors in a more informed manner. For example, Harrower and Brewer (2011)'s popular online tool *ColorBrewer* is very well known in expert visualization communities, and it guides its users to choose appropriate color schemes for their maps and data types while optimizing the suggested color schemes for color vision impaired users (Harrower and Bloch 2006; Harrower and Brewer 2011; Olson and Brewer 1997). Jenny and Kelso (2007)'s *ColorOracle* is similarly a very useful tool to simulate various color deficiencies (Jenny and Kelso 2007). Given that the color vision issues hinder an estimated 8% of the male population as well as 0.5% of the female population (Wong 2011), addressing these issues are important. However, as demonstrated by many other perceptual studies, it is well worth studying the limits of human color perception also with average visual abilities (known as "color-normal" populations; see Asano et al. 2015). For example, humans' ability to discriminate the shades of the same color is limited. In this context, for chromatic discrimination, the *color distance* (the word "distance" here refers to a metric that describes the visual difference between two colors) is a critical factor. Previously, we have empirically validated that increasing the color distance has consistently increased participants' ability to distinguish areas in choropleth and chorochromatic maps (Brychtová and Çöltekin 2015).

In this study, we tackle a lesser-studied perceptual limitation regarding color, that is, how well can we compare (*match* or *discriminate*) colors as the *spatial distance* (gap) between them increases? Studies in perceptual psychology domain on this topic (named "gap effect") appear to be rare, and it is debated whether we fully understand the factors involved in the thresholds of chromatic discrimination as a function of spatial separation (Danilova and Mollon 2006). To the best of our knowledge, however, no empirical tests have been conducted on this question in a cartographic context despite its clear relevance. Furthermore, relevant distance thresholds are not yet identified, that is, how far apart can we place two identical colors and still tell they are identical; or how strong should be the difference (color distance) between them for us to still perceive the difference at various spatial separation levels? Our general hypothesis is that the ability to correctly distinguish both color hue and color value will decrease as the (spatial) distance between two shades increases. To test our hypothesis, we measured performance metrics (can participants effectively and efficiently match or discriminate the colors as we manipulate the distances between them?) in an online study to reach out to a relatively large population sample and variety of monitors; and we conducted a controlled lab study to confirm the performance metrics that was supported by an eye tracking study to better understand participants' visual strategies when comparing the given colors.

Related work

The visual variable "position" and the distance between map symbols

From the perspective of visual variables, by studying the spatial distance and its effect on color discrimination and matching, we investigate the influence of the visual variable *position* and its impact on the usability of color schemes. The visual variable *position* is an interesting variable to consider in most geographic visualizations, because many (though not all) of the features or phenomena attached to a location are not truly *variable* in their positions (i.e., their position is given by their "natural locations" on earth). In traditional mapping sciences, getting the absolute and relative locations of objects and showing their positions on the map correctly is very important for enabling measurements. Cartographic generalization allows some "displacement," but this needs to be rather local and conservative. In other words, distance between map objects is not something we can freely manipulate. Nonetheless, understanding the impact of *position* is important, so that we can manipulate other visual variables and design choices accordingly (e.g., the color distance). If an object of lesser importance is in a prominent position, we can suppress its saliency through other means such as its color or size, or in the opposite case, one can highlight certain features. Besides, certain peripheral elements (such as the legend, interface elements to interact with digital maps) can be designed in relation to our understanding of the impact of *position*. Therefore, *position* can play a substantial role in effectiveness of the map-based decision-making.

The human visual system (HVS) and perceptual factors

Human's limitations in comparing two distant objects in our visual field based on their visual properties such as size or color have strong links to the capacity of our visual system. Human visual field of view covers approximately 100° away from the nose, 60° toward the nose, 75° downward, and 60° upward in two-dimensional space (e.g., Çöltekin 2006). It is important to remember that our perception is not uniform throughout the visual field; we perceive much more detail and see more colors in the center of our vision ("foveal" and "parafoveal") where our contrast sensitivity as well as perception of detail and color decline dramatically toward the periphery, both in two-dimensional (lateral) and three-dimensional (medial) space (Çöltekin 2009; Gordon and Abramov 1977). Furthermore, most of the "meaningful" visual perception is obtained only from a part of the visual field, and this depends largely on the task type (Holmqvist et al. 2011). According to Holmqvist et al. (2011), for example, when viewing photographs of natural scenes, about 10° of the visual field transmits meaningful perceptual signal. For reading, we primarily seem to use foveal perception alone, that is, we perceive letters approximately 3° in the direction of reading and barely 1° at the opposite direction, and text that is located just beyond parafoveal regions in the visual field (i.e., more than 10° horizontally away from the point of view) is not readable (Holmqvist et al. 2011). On the other hand, most importantly for our research, it is suggested that despite the declining color perception, color shades can be distinguished by the peripheral vision to some degree, even if the object is located at the edge of the field of view (Ishiguro and Rekimoto 2011). This is interesting for studying the thresholds of color distance and the spatial gap. In our case, it has relevance especially for eye movement analysis, as it suggests that the objects that are far apart in the visual field might be processed by the HVS partly through mechanisms of the peripheral vision.

While color perception literature is vast and a review of all factors would be beyond the scope of this paper, it is important to note that literature suggests that for color matching, the size of the objects (e.g., colored areas) is an important moderating factor; when the target size is increased, saturation increases, and this impacts what we can see (Gordon and Abramov 1977; Stone 2012; Stone, Szafir, and Setlur 2014; Withouck, Smet, and Hanselaer 2015). This information is important for our study as it informs our experimental design, that is, we control for the size of the areas that we compare.

Yet another relevant aspect of visual perception to our study is that conscious perception and recognition of object properties is affected by other objects in their surroundings (Whitney and Levi 2011). For example, on a visually cluttered map, symbols may be difficult to read because of the presence of other symbols (Phillips and Noyes 1982). This *visual clutter*, therefore, may further reduce our ability to efficiently perform visual search, or distinguish relevant information (Rosenholtz, Yuanzhen, and Nakano 2007), inserting extra demands on the working memory. Finding a target among many distractors in the presence of visual clutter has been studied widely (e.g., Nagy, Sanchez, and Hughes 1990; Wolfe 1994; Rosenholtz, Yuanzhen, and Nakano 2007). In cartography, this "visual search among distractors" occurs naturally when we search for a specific map symbol among other, perhaps somewhat similar, symbols (e.g., similar colors). Similarly to the object size, we control for visual clutter in the experimental design. Furthermore, in geographic visualizations it is common to compare and rank map symbols, which is essentially more complex than visual search (Knapp 1995). As opposed to visual search alone, comparison tasks require not only the perceptual processes that involve locating the target object, but also cognitive tasks such as understanding the meaning of the object and making a decision about it in relation with another. Even though our study is set up to be purely perceptual (i.e., no direct cognitive tasks), it has an implicit ranking task and we believe our results will contribute to the research efforts on understanding visual complexity of maps and effective color use in cartographic design.

Methods

This study examines the participants' ability to correctly distinguish and compare colors on a thematic map with controlled spatial distances between them. Specifically, we study the effect of spatial distance on choropleth and chorochromatic maps colored with sequential and qualitative schemes (respectively). We expect that increasing the gap will lead to decreasing ability to distinguish visualized information. This study was performed in two phases: The first stage was conducted online, thus we will refer to it as the *Web Survey* (the WS); the second part was carried out under controlled laboratory conditions with eye-tracking, thus we will refer to it as "the ET" for the remainder of the manuscript.

Participants

A total of 211 volunteers (121 females, 90 males, 57.4% and 42.6% respectively) have participated in the online experiment (the WS). Seven percent of the participants were in the 16–19-year age bracket, 70% in 20–30, 15% in 31–40, 4% in 41–50, 3% in 51–60, and only 1% of the participants were more than 70 years old. The controlled lab experiment (the ET) had 32 voluntary participants (19 females, 13 males, 59.4% and 40.6% respectively). Age groups in the ET were 93% of 20–30, 4% of 50–60, and 2% of 61–70 years. In both experiments, participants were asked to provide a self-evaluation of their expertise levels in cartography and GIS. In the WS, we obtained data from 54 male and 44 female experts (25.6% and 20.8% respectively, total 46.4% experts), and 36 male and 77 female novices (17% and 36.5% respectively, total 53.6% novices). Proportion of the ET participants was rather similar: 9 male and 8 female experts (28.2% and 25% respectively, total 53.1% experts) and 4 male and 11 female (12.5% and 34.4% respectively, total 46.9% novices) novices. Responses from 15 of the WS participants (10 females, 5 males, 4.7% and 2.3% respectively, total 7.1%), who reported they have some kind of color vision deficiency or have not passed the color vision test with pseudoisochromatic plates have been excluded from the evaluation. None of the participants in the ET experiment had reported color vision deficiency, nor tested negative with the color vision test with pseudoisochromatic plates.

Design and procedure

In this study, we utilized a *color distance* metric that allows quantifying the perceived difference between two colors (CIE 2014). According to Pele and Werman (2012), the most commonly used computational color distance methods are CIE79 and CIEDE2000. The first method calculates the Euclidean distance of the two colors in CIELab color space, while the second one, based on the same color space, contains compensation for neutral colors, lightness, chroma, and hue to reach higher perceptual uniformity. Based on a literature review (e.g., Carter and Huertas 2009; Yang, Ming, and Yu 2012), we decided to apply CIEDE2000 to calculate the color difference between shades of sequential and qualitative color schemes (for the mathematical specifications of CIEDE2000, see Sharma, Wu, and Dalal 2005). This method was shown to be suitable for calculations of both small and large color distances (Carter and Huertas 2009).

Both the WS and the ET were designed as within-subject factorial experiments with randomized stimuli (for a definition of these experimental terms, see e.g., Rubin and Chisnell 2008). Firstly, we executed and analyzed the WS. The main independent variable in this study was the stimuli with controlled *spatial distances* (see Section 3.3). For both experiments, our primary dependent variables were *accuracy* and *response time*. However, in the ET, we also used eye movement metrics *fixation frequency, fixation duration, scanpath speed,* and a *gaze transition analysis*. No time limits were imposed in either study. During the controlled experiment (the ET), we included a training task prior to the experiment to ensure that participants fully understand the question and they were allowed to ask the experimenter for clarifications during the training session (this was not possible for the WS for practical reasons).

We presented two *purely perceptual* tasks to the participants (see Figure 1):

(a) "Locate two areas on the map that are marked with a dot, compare them, and decide whether these areas are of the same color." Participants could choose one of these three answers: (1) yes, marked areas are of the same color; (2) no, marked areas are not of the same color; and (3) I don't know, I am not able to tell whether these colors are the same or not. This task was used for both sequential and qualitative color schemes in the ET and the WS.

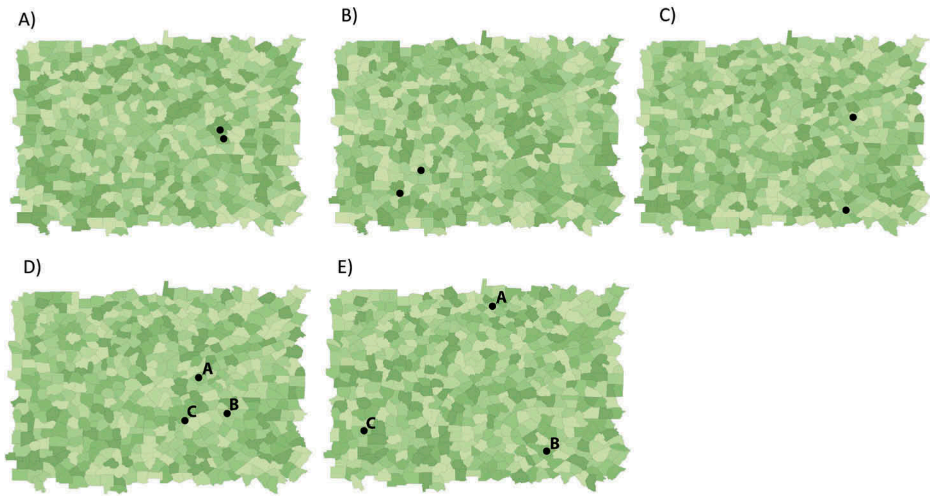


Figure 1. Demonstration of experimental stimuli. Comparing two areas (A) Δd≈small, (B) Δd≈medium, and (C) Δd≈large on maps with sequential schemes (qualitative schemes were studied in a same way). Comparing three areas: (D) Δd≈medium and (E) Δd≈large. Markings (dots and letters) were exaggerated on article illustrations for better readability.

(b) "Locate three areas on the map that are marked with a dot and each labeled with a letter, and arrange them from lightest to darkest shade." This task was used only for sequential color schemes and only in the ET.

Materials

The WS was implemented in the open source application *LimeSurvey* (The LimeSurvey project 2011) and participants were asked to keep the conditions "as usual," that is, under which they normally work with the computer (to obtain some degree of ecological validity). They were also asked *not* to manipulate their screen settings and the room lighting during the survey to obtain some degree of control during the session. The WS stimuli had embedded sRGB ICC profile (sets of information necessary to convert color data between native device color spaces and device-independent color spaces (ICC 2014) and their size was 800 × 600 px.

The ET study was carried out under controlled conditions in the laboratory at Palacký University Olomouc, equipped with a low-frequency contactless eye tracker *SMI RED 250* (SensoMotoric Instruments 2013) with a sampling frequency 120 Hz. Stimuli were projected on 23′ LG Flatron monitor IPS231P. ET stimulus size was 1920 × 1080 px. Stimuli had embedded ICC profile of the laboratory monitor. Experiment was prepared and presented in *SMI Experiment Center*™. Fixation detection was performed through the *SMI BeGaze*™ using ID-T (dispersion threshold algorithm). Dispersion threshold was set to 50 px and a minimum length of 80 ms (suggested by Popelka 2014). Calculation of basic eye-tracking metrics and areas of interest (AOI) transitions has been performed in *OGAMA* (Voßkühler 2013) and statistical analysis of the data in the statistical software *R* (R Core Team 2013).

As visual stimuli, we created maps that represent fictional territories with areas of approximately equal size (30–45 px at WS; 60–90 px at ET) for experimental control purposes. Additionally, map content and geometry was significantly simplified to reduce the influence of possible confounding variables.

The color distances between the selected areas were controlled to be $\Delta E_{00} = 2$, $\Delta E_{00} = 4$, $\Delta E_{00} = 6$, $\Delta E_{00} = 8$, $\Delta E_{00} = 10$, or $\Delta E_{00} = 0$ (same color). Spatial distribution of other colors was randomized to avoid (or distribute) possible simultaneous contrast effect as much as possible, even though this effect is not easily eliminated (Bláha and Štěrba 2014). We selected six shades of green for sequential color schemes and six color hues with constant lightness (yellow, orange, red, violet, blue, and green) for qualitative color schemes (Table 1 and Figure 1). In case of the sequential schemes, we kept values a and b stable, while alternating value L (lightness) of a specific amount to reach the desired color distance. Qualitative schemes were designed so the lightness was kept approximately stable, while variations of a and b resulted in desired hue of shades. Our choice of examined colors was driven rather for pragmatic reasons; we were not able to test more colors for both sequential and qualitative schemes, simply because the experiment would be unfeasibly long.

Table 1. Color codes of (A) sequential and (B) qualitative color scheme shades.

ΔE_{00}	class	L	a	b	R	G	B
(A)							
10	A	94,80	−30,00	30,00	201	255	179
	B	79,45	−30,00	30,00	159	211	138
	C	66,10	−30,00	30,00	123	174	104
	D	54,67	−30,00	30,00	93	143	76
	E	44,67	−30,00	30,00	68	117	52
	F	33,12	−30,00	30,00	39	89	26
8	A	94,80	−30,00	30,00	201	255	179
	B	82,39	−30,00	30,00	167	219	146
	C	71,22	−30,00	30,00	136	188	117
	D	61,34	−30,00	30,00	110	161	92
	E	52,63	−30,00	30,00	88	138	71
	F	44,58	−30,00	30,00	67	117	52
6	A	54,71	−30,00	30,00	93	143	76
	B	61,34	−30,00	30,00	110	161	92
	C	68,63	−30,00	30,00	129	181	110
	D	76,63	−30,00	30,00	151	203	131
	E	85,40	−30,00	30,00	175	228	154
	F	94,80	−30,00	30,00	201	255	179
4	A	66,15	−30,00	30,00	123	174	104
	B	71,24	−30,00	30,00	136	188	117
	C	76,65	−30,00	30,00	151	203	131
	D	82,40	−30,00	30,00	167	219	146
	E	88,51	−30,00	30,00	184	237	162
	F	94,80	−30,00	30,00	201	255	179
2	A	79,47	−30,00	30,00	159	211	138
	B	82,39	−30,00	30,00	167	219	146
	C	85,40	−30,00	30,00	175	228	154
	D	88,50	−30,00	30,00	183	237	162
	E	91,70	−30,00	30,00	192	246	171
	F	94,80	−30,00	30,00	201	255	179
(B)							
10	A	94,70	7,06	1,24	255	235	236
	B	94,80	3,00	11,00	255	238	217
	C	99,20	−4,70	13,93	255	255	224
	D	95,00	−7,65	1,50	226	245	236
	E	94,00	−5,40	−10,00	216	242	255
	F	94,00	2,60	−9,20	234	237	253
8	A	95,50	5,65	0,99	254	239	239
	B	95,90	2,20	8,75	255	241	225
	C	99,30	−3,76	11,03	255	255	230
	D	95,00	−5,28	1,20	231	244	236
	E	95,00	−4,10	−8,00	224	244	254
	F	95,00	2,10	−7,36	238	240	253
6	A	95,00	3,80	0,74	249	238	237
	B	96,00	1,60	6,45	252	242	229
	C	99,50	−2,47	7,18	255	255	238
	D	95,00	−3,61	0,92	234	243	237
	E	95,00	−2,80	−5,91	229	243	250
	F	95,00	1,70	−5,50	239	240	249
4	A	95,00	2,39	0,50	246	239	238
	B	96,00	1,00	4,02	249	243	234
	C	99,70	−1,30	3,81	255	255	245
	D	95,00	−2,01	0,62	237	242	238
	E	95,00	−1,70	−3,91	233	242	246
	F	95,00	1,20	−3,66	240	240	246
2	A	95,00	−1,41	0,25	238	242	238
	B	95,00	−0,60	2,20	241	241	235
	C	99,00	−0,45	1,41	252	252	247
	D	97,00	0,80	0,30	248	246	244
	E	95,00	0,80	−1,90	241	240	242
	F	95,00	−0,60	−1,84	238	241	242

Theoretically, color distance metric should report the visual difference equally well regardless the hue or lightness differences, that is, if we compare two color pairs – one distinguished with hue, the other with lightness – both can be of the same color distance (numerically).

We used 106 stimuli in the WS and 53 in the ET experiment. Stimuli used in the ET are a subset of the ones from the WS, except 12 of them that present a different question type (compare three areas). On each stimulus, two or three areas (depending on the task) were selected and marked with a dot (Figure 2). The spatial distribution of these areas was controlled according to the following criteria: (a) next to each other ($\Delta d \approx$small, note that only in the WS), (b) at mid-distance; that is, in between the 2 or 3 depicted areas lie 2 or 3 other areas ($\Delta d \approx$ medium), and (c) at two extremes; that is, in between the 2 or 3 depicted area lie 8 to 13 ($\Delta d \approx$ large). Corresponding visual angles for these distances were calculated by trigonometric function of the distance of two compared points (on the stimulus) and the distance of participant's eye from the monitor. Since this distance could not be controlled in the WS, we estimated an average 50 cm as a possible sensible average. This estimation is based on informed reasoning as we decided to fix the size of the stimuli to 800 × 600 pixels for possible broad range of displays participants would use. For this stimuli size, our pilot testing suggested that 50 cm may be a reasonable assumption. Based on this, in the controlled lab study where we had a larger display, participants were seated, approximately 70 cm away from the monitor to obtain the same visual experience, that is, we used 70 cm as the input in our calculations of the visual angle. Dimensions of stimuli and viewing angles are described in Table 2.

We conducted pilot sessions for the ET. Our observations in these pilot sessions revealed that exploring $\Delta d \approx$small based on eye movements does not offer much added value over what we measured in the WS. We observed that the gaze position remained more or less on the border of the compared areas for $\Delta d \approx$medium condition (Figure 3) and due to the accuracy limits of the eye tracker, it would not be possible to exactly determine which area was visited at that moment. We believe this may be caused by the HVS's ability to compare close areas using parafoveal or peripheral vision; that is, without necessarily moving the eyes to get a better foveal vision (e.g., Danilova and Mollon 2010).

Results

We first analyzed the standard performance metrics *accuracy* and *response time*, and afterwards selected *eye-tracking metrics* together with *gaze transition* matrices. Accuracy and response time measurements were mainly evaluated using the larger sample acquired

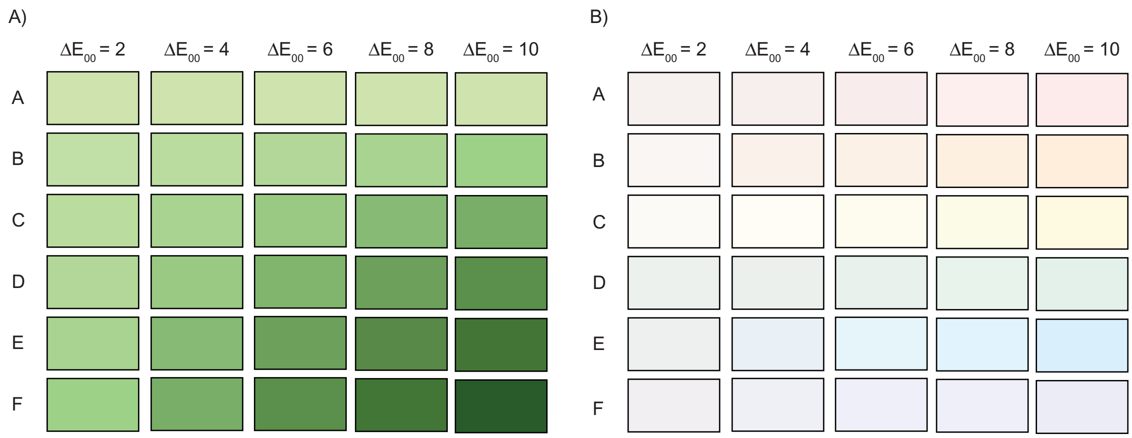


Figure 2. Tentative view of examined color (A) sequential and (B) qualitative schemes.

Table 2. Spatial distance level for WS and ET with corresponding visual angles.

Spatial distance	**The WS study** estimated observing distance 50 cm, estimated pixel size 0.27 mm, and stimulus dimension 800 × 600 px	**The ET study** observing distance 70 cm 24″ screen, pixel size 0.27 mm, and stimulus dimension 1600 × 1200 px
Δd ≈ small	Less than 1.5°	not applied
Δd ≈ medium	2.5–3.0°	2.5–3.5°
Δd ≈ large	6.0–7.5°	10.0–13.0°

from the WS. Where possible, these were further verified with performance metrics in the controlled experiment (ET), which had a subset of experimental conditions and a smaller population sample. We then conducted an eye movement analysis where we studied *fixation frequency, fixation duration, scanpath speed,* and a *gaze transition analysis*. We hypothesized the impact of spatial distance on our measures as follows:

(H1) Larger spatial distances will lead to overall *lower accuracy* (more mistakes) and *longer response times*

(H2) Larger spatial distances will cause *longer fixation durations, higher fixation frequencies, longer scanpaths*, and an *increasing number of revisits* (transitions) between compared areas

(H3) Sequential and qualitative color schemes will show no difference in the accuracy, response time, or eye-tracking metrics at the same spatial distance

Results are reported separately based on the complexity of the experimental tasks (two or three areas to compare).

Compare two areas: accuracy

We summed and analyzed the observed accuracy in groups based on spatial distance, while keeping the sequential and qualitative color schemes separate. We coded responses as "unsuccessful" (i.e., inaccurate) if participants marked two colors identical where they were not, and vice versa. We also considered "I don't

Figure 3. Visualization of raw ET data (left) and fixations (right) of one selected participant from the pilot study while comparing two areas with Δd≈small.

know" responses unsuccessful, because if people could not tell if the examined colors were identical or different, they could not discriminate the given color pair.

Chi-square goodness of fit test showed that the accuracy in the WS is dependent on the spatial distance for both color scheme types: sequential (χ^2 = 405.31, df = 2 and p < .01) also qualitative (χ^2 = 65.57, df = 2 and p < .01). The ET results showed this only with sequential schemes (χ^2 = 22.16, df = 1 and p < .01). Qualitative schemes did not yield significant difference between Δd≈medium and Δd≈large (χ^2 = 1.89, df = 1 and p = 0.17). Therefore, the first hypothesis (H1) was confirmed in the WS: Participants had more trouble comparing both quality and quantity of the two areas when the gap between them was larger (Figure 4a). We could not, however, confirm this finding in the controlled study; that is, in the ET, we observed no significant decrease in accuracy with increased spatial distance for this "compare two areas" condition (Figure 4b). The differences in the findings between the WS and the ET suggest that spatial distance has bigger impact on distinguishing the quantities (shades of sequential schemes) rather than qualities (different hues), thus we reject the hypothesis H3. Nonetheless, note that the number of participants was considerably higher in the WS, and if we chose to interpret the descriptive statistics, we also see a small difference in the ET (Figure 4b).

Compare two areas: response time

Overall, in the WS, questions were answered remarkably slower than in the ET (Mdn_{WS} = 6.95 s, Mdn_{ET} = 3.05 s). We believe this vast difference is mainly due to the differences in the experiment procedure: For the WS, the response time includes task solving *and* marking the response (they were displayed on the same page). For the ET, the answer form was displayed separately from the stimulus, and only the viewing time was recorded (not the time spent marking the response). Nonetheless, to avoid possible bias from the interruptions in the WS (participants could check email, answer calls, have a coffee, etc.), trials that took longer than 90 seconds (based on the maximum response time observed in the ET), that is, 65 outliers, were removed from the analysis for the WS. We had a total of 23,366 trials (211 participants × 106 conditions), and we have removed 65 of them, which were distributed over several conditions; thus we believe this operation has no significance for the repeated measures analysis.

Response times for the WS were further analyzed with Kruskal–Wallis (H) and Mann–Whitney (U) test based on correct answers only. Kruskal–Wallis test on WS data revealed that response time significantly increases with increasing spatial distance for both color scheme types (sequential: H = 369.74, p < .01; qualitative: H = 32.53, p < .01), Table 3. This confirms the hypothesis (H1). For the ET data, Mann–Whitney test showed significant difference for qualitative schemes only (Table 4).

Comparison between sequential versus qualitative color schemes with Mann–Whitney test indicates that for the same spatial distance (and same color distance), sequential schemes were harder to match than the qualitative ones: the median response time at Δd≈medium and Δd≈large was significantly longer with sequential schemes during both the WS and ET (Tables 3 and 4). Similarly as the accuracy results, this finding is in conflict with hypothesis (H3), strongly suggesting that for better discriminability, the sequential color schemes should be designed more carefully regarding the spatial distance between colors.

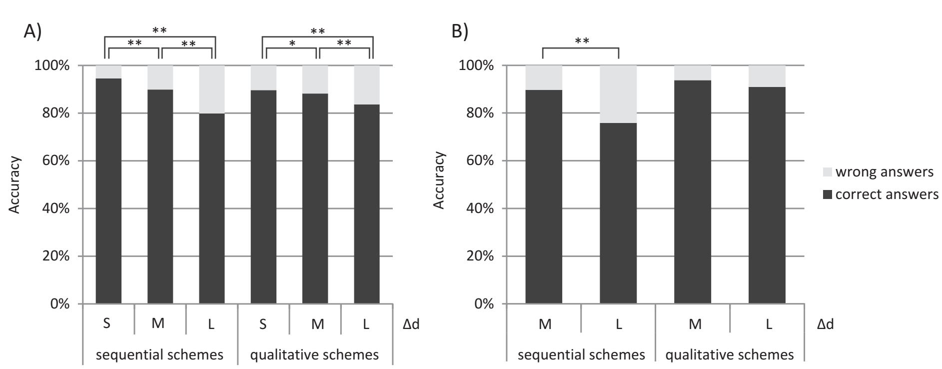


Figure 4. The accuracy rates for the WS (A) and the ET (B). (S, M, and L refer to small, medium, and large spatial distance respectively).[1.]

Table 3. Median values of observed response time for spatial distance levels and color scheme types.

Color scheme	Δd	Response time WS [s]	Response time ET [s]
Sequential	all	6.95	3.05
	small	6.40	–
	medium	6.92	2.81
	large	7.82	3.16
Qualitative	all	6.62	2.26
	small	6.47	–
	medium	6.67	2.13
	large	6.80	2.49

Table 4. Results of Mann–Whitney (U) and Kruskal–Wallis (H) tests comparing the response times between spatial distance levels (S-small, M-medium, and L-large) and color scheme types. Significant differences are in bold.

	The WS		The ET	
Condition	U (H)	*p*	U	*p*
Sequential ~ Qualitative	**44497742**	**<0.01**	176059.5	0.46
Sequential S ~ M ~ L	**(H) 369.74**	**<0.01**	–	–
Qualitative S ~ M ~ L	**(H) 32.53**	**<0.01**	–	–
Sequential M ~ L	–	–	36277.5	0.28
Qualitative M ~ L	–	–	**43444.5**	**0.04**
Sequential S ~ Qualitative S	6183307	0.41	–	–
Sequential M ~ Qualitative M	**4733845**	**<0.01**	**52634.0**	**<0.01**
Sequential L ~ Qualitative L	**3937946**	**<0.01**	**51011.5**	**<0.01**

Compare two areas: analyses of eye-tracking data

During the ET experiment, only Δd≈medium and Δd≈large were evaluated (see section *Materials*). The influence of examined condition was further evaluated by analyses of selected eye-tracking metrics: fixation frequency, average fixation duration, and scanpath speed. The meaning of these metrics can be interpreted as follows: more overall fixations [count/s] indicate less searching efficiency (Goldberg and Kotval 1999). Longer fixation duration [ms] may signify increased difficulty in understanding the meaning of information, or that the fixed object is in some sense more interesting or relevant to the task (Poole and Ball 2005). Lower scanpath speed [viewed px/s] can be interpreted as a certain level of (self) confidence and more careful deliberation while conducting visual search (Brychtová and Çöltekin 2014).

The hypothesis (H2) was confirmed at both color scheme types with average *fixation duration* and *scanpath speed*. Mann–Whitney test confirmed significantly shorter fixations and faster scanpath for Δd≈large (Tables 5 and 6). On the other hand, the *fixation frequency* was significantly different only at qualitative schemes. The hypothesis (H3) was supported by all examined metrics: there was no significant difference between sequential and qualitative schemes at the same spatial distance (Tables 5 and 6).

Table 5. Median values of eye-tracking metrics based on spatial distances Δd≈medium (M) and Δd≈large (L) and color scheme types sequential and qualitative.

Color scheme	Δd	Fixation frequency [count/s]	Average fixation duration [ms]	Scanpath speed [px/s]
Sequential	M and L	4.44	177.95	588.60
Qualitative	M and L	4.42	178.10	572.11
Sequential	M	4.39	181.7	397.71
	L	4.48	175.0	890.03
Qualitative	M	4.28	187.85	378.76
	L	4.51	168.50	864.19

Table 6. Results of Mann–Whitney: differences of eye-tracking metrics based on spatial distances Δd≈medium (M) and Δd≈large (L) and color scheme types (sequential and qualitative). Significant differences are in bold.

	Fixation frequency		Average fixation duration		Scanpath speed	
Condition	U	*p*	U	*p*	U	*p*
Sequential ~ Qualitative	176059.5	0.46	**57012.5**	**<0.01**	**6436.5**	**<0.01**
Sequential M ~ L	36063.5	0.23	**43902.0**	**<0.01**	**6848.0**	**<0.01**
Qualitative M ~ L	**42142.0**	**<0.01**	**57012.5**	**<0.01**	**6436.5**	**<0.01**
Sequential M ~ Qualitative L	45657.5	0.21	41790.5	0.54	45900.5	0.17
Sequential M ~ Qualitative L	42269.5	0.83	43614.5	0.66	43531.0	0.69

In the next step, we performed a gaze transition analysis between AOIs. Transitions are defined as movements from one AOI to another in the eye movement literature. Their interpretation depends on the concrete task (Holmqvist et al. 2011). In our case, we expect that gaze transitions would be straight back and forth between the depicted areas in order to compare their colors. We interpreted that the more transitions participants exhibited, the more troubles they had in solving the task. On each stimulus, we defined two AOIs covering the depicted area and its direct surroundings (A and B) and one AOI covering the rest of the map (X).

Number of transitions was averaged with number of participants and stimuli for each examined condition. Number of transition A↔B at Δd≈medium was unexpectedly much higher than at Δd≈large, while this difference is stronger on sequential schemes. On the contrary, number of transition A↔X and B↔X is higher at Δd≈large (Figure 5). In a deeper investigation of the raw ET data visualization, we found that on Δd≈medium participants are able to "jump" directly from A to B, while at Δd≈large they usually do "inter-fixation" in the AOI X. Overall average fixation

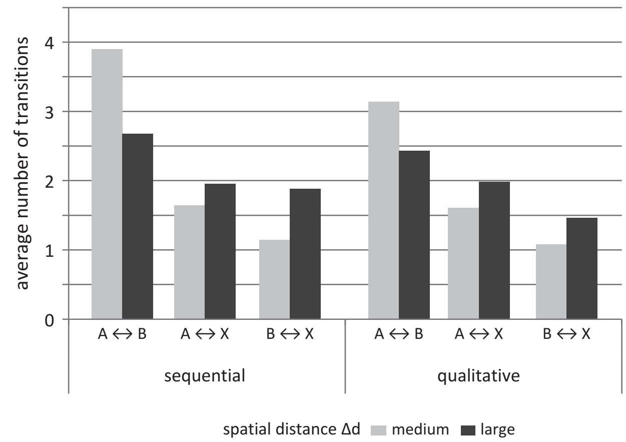


Figure 5. Averaged number of transition between AOIs for one participant and one stimulus as an example.

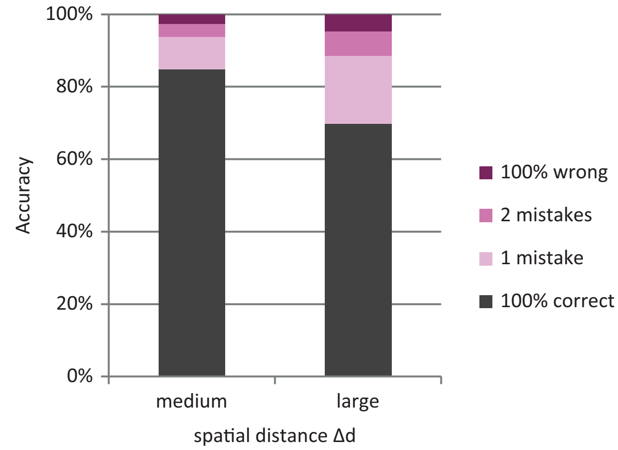


Figure 6. Accuracy of answers while comparing the colors of three areas.

durations registered in X is 221 ms, while in A and B it is 301 ms and 303 ms, respectively. This suggests that short-lasting fixations in X might serve as a kind of "layover" and not to read the information. Another reason for transitions from main AOIs to X is possibly the calibration accuracy of the eye tracker. Any instability in the calibration can lead to recordings in which the fixations do not reflect the true position of the gaze, thus a gaze that was intended at the relevant AOI may be counted outside of it (see the Appendix 1). At Δd≈medium there were, in average, 0.8 more transitions between A↔B on sequential than qualitative color schemes, which indicates trouble when comparing shades of the same color (green, in our test), while individual hues appear to be easier to distinguish.

Compare three areas: accuracy

Comparing and ordering three areas according to the visualized quantity from the smallest (lightest color) to the highest (darkest color) proved to be generally more difficult than the previous task (compare two areas). This task was designed only to evaluate sequential schemes (as ordering them would not make sense for qualitative schemes under studied conditions). The accuracy was assessed with regard to the number of errors. An answer was considered 100% correct when ordering was right for all the three areas. Ordering task allowed marking two colors identical (this was sometimes the case); thus participants could make one, two, or three mistakes. For example, if the correct answer is A < B < C[2]; then A = B < C is coded as "1 mistake," B < A < C is coded as "2 mistakes," and C < A < B is coded as 100% wrong. Shorter distances between compared areas yielded in more accurate answers (χ^2 = 12.58, df = 1, p < .01): 85% of fully correct answers were at Δd≈medium, while only 69% were at Δd≈large (Figure 6). Thus, the hypothesis (H1) was confirmed.

Compare three areas: response time and eye-tracking metrics

Differences between Δd≈medium and Δd≈large were significant for *response time* (Mdn_{medium} = *5.05 s*, Mdn_{large} = *6.92 s*, W = 7266, p < .01), *average fixation duration* (Mdn_{medium} = *179.60 s*, Mdn_{large} = *165.05 s*, W = 13223, p < .01), and *scanpath speed* (Mdn_{medium} = *428.61 s*, Mdn_{large} = *1118.87 s*, W = 1986.5, p < .01). According to these metrics, the more problematic spatial distance was Δd≈large (Figure 9), thus, hypotheses (H1) and (H2) were confirmed. Fixation frequencies did not yield significant differences similarly as in the task "compare two areas."

Areas of interest were defined in the same manner as in the previous task: AOIs A, B, and C were assigned to the depicted areas (A corresponds to the lightest and C to the darkest color). The rest of the stimulus was marked as a single AOI as X. Overall, we observed more transitions between AOIs of closer colors (A↔B and B↔C), while between A↔C (more distinguishable as their color distance ΔE_{00} is bigger), less transitions were recorded. In this task, again, more transitions were observed at the shorter spatial distance Δd≈medium. At Δd≈large, there is even less direct transitions between main AOIs (A, B, and C), while distinctly more transitions are between A, B, or C and X (Figure 7). Average fixation frequency in AOIs was 270 ms (A), 269 ms (B), 269 ms (C), and 195 ms (X).

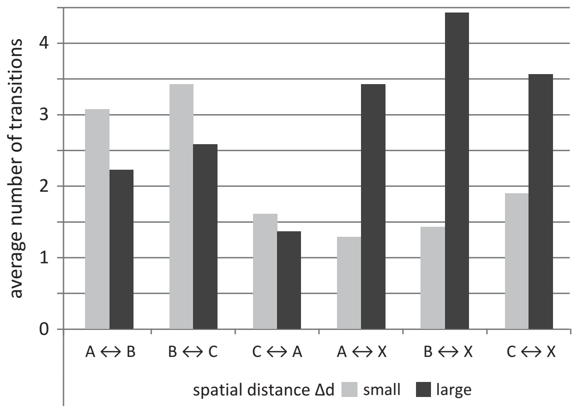


Figure 7. Averaged number of transitions.

The interactions between spatial distance and color distance

In the last phase, we examined interactions between the spatial distance (Δd) and the color distance (ΔE_{00}). In other words, we wanted to find out if, for the same color distance, it is harder to match and compare colors when they were spatially more separated. This evaluation is based only on the *accuracy* and *response time* from the WS. Our results regarding *accuracy* analysis suggest that, overall, participants' ability to distinguish between the two areas of sequential schemes declines when their mutual spatial distance grows (Figure 8a). This is particularly clear at color distances ΔE_{00} = 2, 4, and 6. Accuracy at ΔE_{00} = 10 is particularly high (more than 90%) at all spatial distances.

Qualitative color schemes did not yield a strong variability in accuracy across spatial distance levels (Figure 8b). At all color distances (except ΔE_{00} = 2); participants were most accurate when comparing neighboring areas (Δd≈small). Accuracy rates at Δd≈medium and Δd≈large were approximately equal, which suggests that comparing symbols based on color *hue* is not as vulnerable to the spatial distance between compared areas. Differences of accuracy between all spatial distance levels are not significant at ΔE_{00} = 10. We interpret these results as such that perceptual difference between shades of qualitative schemes should be at least ΔE_{00} = 10, because this level of color distance allows to correctly distinguish map symbols even if they are far apart. This is partially true also for the sequential schemes, that is, ΔE_{00} = 10 is also the "safest" color distance among the ones we tested.

The *response time* analysis confirmed that sequential color schemes indeed have a higher sensitivity to spatial distance: at all color distances (ΔE_{00}), the shortest time was observed for Δd≈small and the longest for Δd≈large (Figure 9a). This trend did not occur with qualitative schemes (Figure 9b).

Discussion

The goal of the presented study was to investigate the effect of spatial distance between map symbols on the discriminability of colors. In a two-stage user experiment, participants executed two perceptual tasks varying in their level of complexity: decide if two areas on the given choropleth (or chorochromatic) map are of the same color, and rank three areas according to their color value. We controlled the spatial distance between the compared areas at three levels (Δd≈small, Δd≈medium, and Δd≈large), and manipulated the design of stimuli so they could be colored with sequential and qualitative schemes with equal color distances between their intervals/classes (ΔE_{00} = 0, ΔE_{00} = 2, ΔE_{00} = 4, ΔE_{00} = 6, ΔE_{00} = 8, and ΔE_{00} = 10). We measured and statistically analyzed traditional performance metrics (accuracy and response time), and selected eye-tracking metrics (fixation frequency, fixation duration and scanpath speed) as well as transitions between AOIs. Based on the analysis of all metrics, we found that, overall, increasing the spatial distance between colors has a consistent negative impact on the ability to differentiate them with both sequential and qualitative schemes.

A finer analysis on accuracy, response time, and AOI transitions further suggested that at the same ΔE_{00}, sequential schemes (shades of the same color) are more difficult to distinguish than qualitative schemes (different hues). In other words, at the same spatial distance *and* the same color distance, discriminating colors with qualitative color schemes was overall easier for the participants. Most likely explanation for this difference comes from a feature of human cognition: earlier literature demonstrated that people can better distinguish colors that they are able to name (Brewer 1996; Brown, Lindsey, and Guckes 2011; Lenneberg 1961; Özgen 2004), possibly exploiting verbal memory in addition to the visual and spatial memories. Our selection of examined qualitative colors could have possibly increased the influence of "nameability" in this study (basic rainbow colors, which should be easily nameable for color-normal populations). A different set of qualitative colors might lead to different results, that is, whether color naming dominates or color distance metrics dominate should be further tested.

Theoretically, the CIEDE2000 color distance computational model is formulated as such that one can reasonably expect that our results can be transferred to other colors. If the color distance (ΔE_{00}) is the same

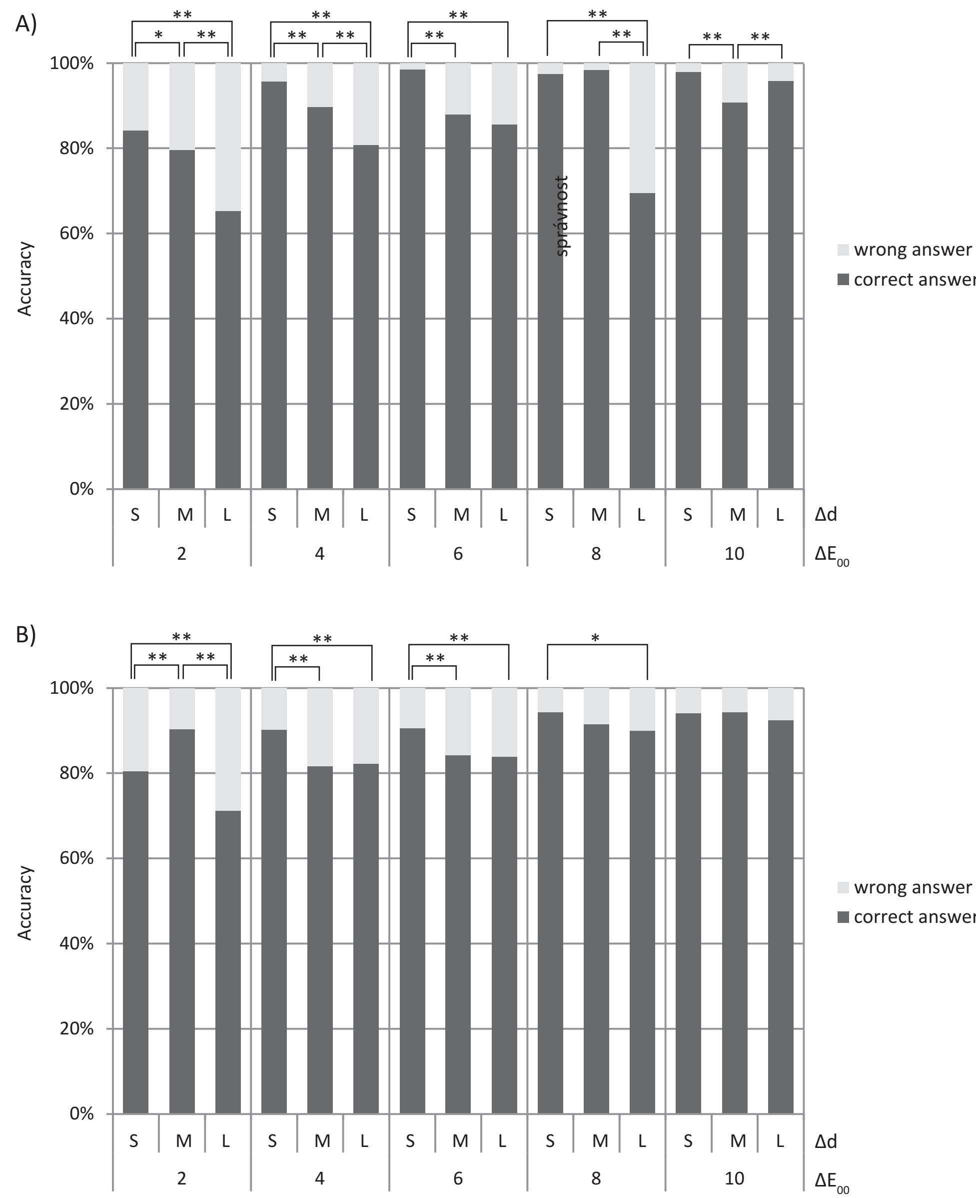


Figure 8. The correlation between spatial (Δd) and color distance (ΔE_{00}): Accuracy of answers for (A) sequential and (B) qualitative color schemes. (S, M, and L refers to small, medium, and large spatial distance respectively). (See endnote 1 for *, **)

between two colors according to the CIEDE2000; regardless their saturation, lightness or hue, the perceptual experience should be the same. Empirically, we validated this hypothesis in another study where we compared green and red sequential schemes and found that the results largely agree (Brychtová and Vondráková 2014). This means that our results, at least to some degree, can be generalized to other colors. However, further empirical experiments are necessary to validate our findings with other colors as the CIELAB exhibits different levels of sensitivity for various hues.

It is also important to remind the reader that we examined only six-class color schemes. The number of classes could affect our results, and this requires further testing. Is it sufficient to keep the color distance ΔE_{00} = 10 for schemes with more than six classes, or should we increase the visual distance of such color schemes to obtain the same level of usability?

Our experimental results do not provide "the final answer" on what is the optimal color distance value (or an index of values for different use cases) yet. However, the awareness that color distance and spatial distance are important factors in map readability could help cartographers to streamline their map-making process. It appears that, so far, color schemes are designed rather by trial and error. Cartographers vary the color difference based on their intuitive judgment. Beginners

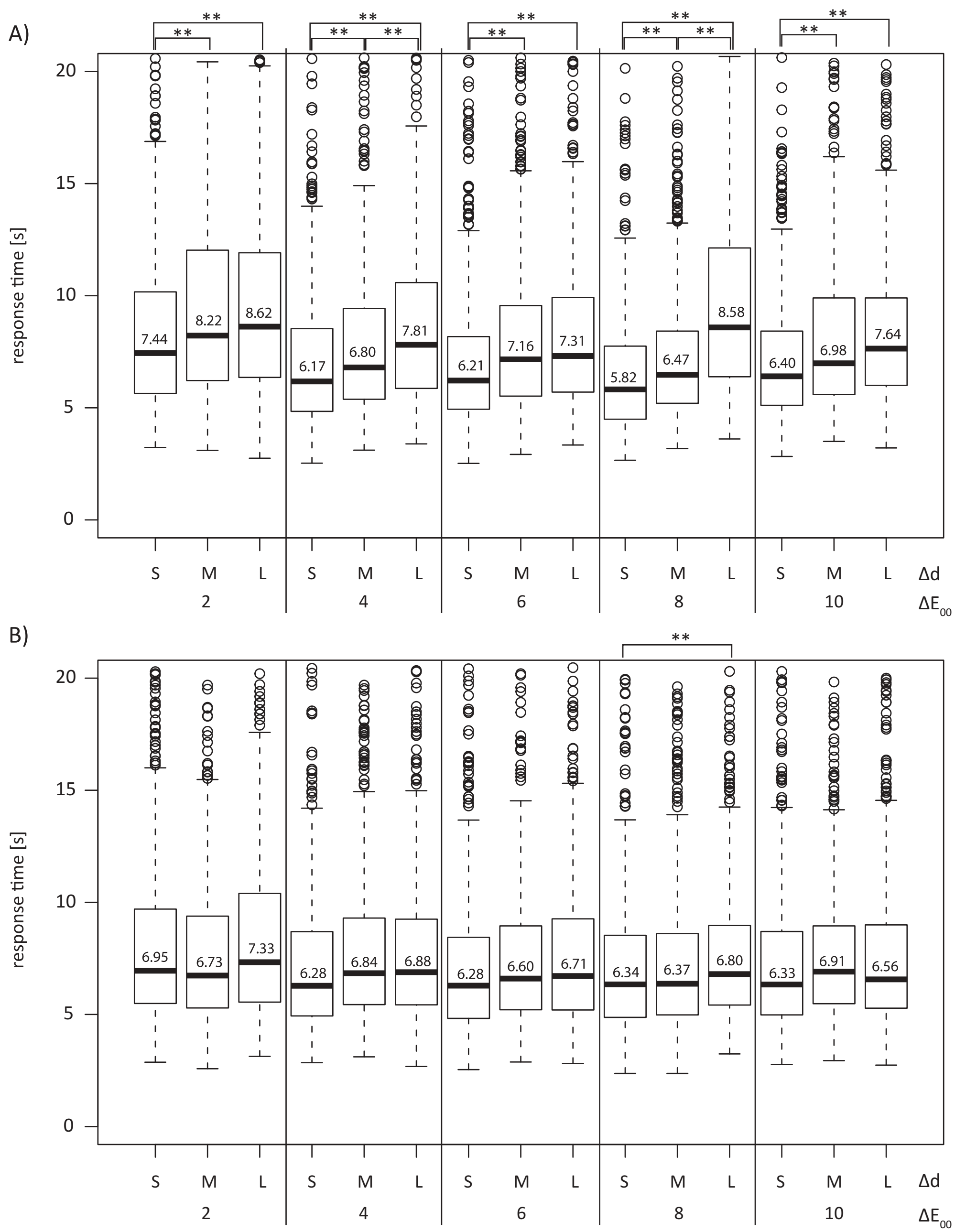


Figure 9. The correlation between spatial (Δd) and color distance (ΔE_{00}): Response time for (A) sequential and (B) qualitative color schemes. (S, M, and L refer to small, medium, and large spatial distance respectively). Median values for each examined condition are given inside boxplots.

make mistakes such as trying to design yellowish sequential schemes with the same (large) amount of classes as blueish ones, without being aware of the dimensions of the color space (see Appendix 2 for a demonstration of various color schemes and their classes distinguished with ΔE_{00} = 10, starting from saturated colors toward white).

Even though in most cases the spatial distribution of map symbols is "given" by the location of mapped features or phenomena, we can manipulate the *color distance* as the spatial distance grows, or we can modify the other visual variables to reduce the perceptual effect. Furthermore, we can be alert about where we place the object symbols, legend, or similar peripheral information that are important for color comparison, and not overlook the fact that what is easy to distinguish at short distances, does not have to be distinguishable across the whole map. Empirically determined color distance values could help us avoid such basic mistakes and save some time from testing new color schemes on a trial-and-error basis.

Conclusions and outlook

Our findings regarding the qualitative and sequential schemes *do not* suggest that we should prefer qualitative color schemes; in fact these two color scheme types are used for different purposes. However, we believe our findings contradict the assumed capabilities of precisely defined colorimetric models, such as CIEDE2000, to predict perceived difference in any context. From a more applied perspective, our results demonstrate that we need to make different design decisions with sequential and qualitative color schemes. Specifically, we can use different ΔE_{00} values for the two color schemes; for qualitative color schemes, smaller color distances allow visually differentiating colors than for sequential color schemes. Our results clearly suggest that for the qualitative and partially sequential colors, color distance $\Delta E_{00} = 10$ yields considerably higher levels of accuracy in color discrimination, even when the spatial gap between the two colors is relatively large; thus we recommend this "more conservative" color distance when designing sequential schemes.

We believe that a set of experiments to understand whether our results are generalizable for displaying different *number of classes* would further our knowledge on this topic. Another interesting direction would be to investigate the interactions between other factors *and* the spatial distance (e.g., the visual clutter, possibly expressed as the number of features/classes displayed over the distance covered, or size of the symbols) and individual differences such as gender or age.

To design color schemes based on the CIEDE2000 color distance computations, we developed a free online software tool "SequentialColorSchemeGenerator" Brychtová and Doležalová (2015), which can be found at http://eyetracking.upol.cz/color. This tool enables the cartographers (as well as visualization designers, GIS users, and others interested in design) to follow our recommendations and adapt the concept of "color distance" in their maps, should they wish to do so.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the ESF under Grant [grant number StatGIS CZ.1.07/2.3.00/20.0170].

Notes

1. We differentiate between varying confidence levels of 0.01 and 0.05 with a notation of a two asterisks (**) and single asterisk (*) respectively. This notation will be used throughout the paper.
2. A < B < C means that A is lighter than B and B is lighter than C; A = B means color shades are identical.

References

Al-Ghamdi, A. M. 2014. "Optimising the Selection of a Number of Choropleth Map Classes." In *Thematic Cartography for the Society*, edited by T. Bandrova, M. Konecny, and S. Zlatanova, 61–77. Basel: Springer International. doi:10.1007/978-3-319-08180-9_6.

Asano, Y., M. D. Fairchild, L. Blondé, and P. Morvan. 2015. "Color Matching Experiment for Highlighting Interobserver Variability." *Color Research & Application*. Advance online publication. doi:10.1002/col.21975.

Bertin, J. 1983. *Semiology of Graphics: Diagrams, Networks, Maps. Components*. Madison: University of Wisconsin Press.

Bjørke, J. T. 1996. "Framework for Entropy-based Map Evaluation." *Cartography and Geographic Information Science* 23 (2): 78–95. doi:10.1559/152304096782562136.

Bláha, J. D., and Z. Štěrba. 2014. "Colour Contrast in Cartographic Works Using the Principles of Johannes Itten." *The Cartographic Journal* 51 (3): 203–213. doi:10.1179/1743277414Y.0000000084.

Brewer, C. A. 1994. "Color Use Guidelines for Mapping and Visualization." In *Visualization in Modern Cartography*, edited by A. M. MacEachren and D.R.F. Taylor, 123–147. Tarrytown, NY: Elsevier Science.

Brewer, C. A. 1996. "Guidelines for Selecting Colors for Diverging Schemes on Maps." *The Cartographic Journal* 33: 79–86. doi:10.1179/caj.1996.33.2.79.

Brewer, C. A., A. M. MacEachren, L. W. Pickle, and D. Herrmann. 1997. "Mapping Mortality: Evaluating Color Schemes for Choropleth Maps." *Annals of the Association of American Geographers* 87 (3): 411–438. doi:10.1111/1467-8306.00061.

Brown, A. M., D. T. Lindsey, and K. M. Guckes. 2011. "Color Names, Color Categories, and Color-Cued Visual Search : Sometimes, Color Perception Is Not Categorical." *Journal of Vision* 11 (12): 1–21. doi:10.1167/11.12.2.

Brychtová, A., and A. Çöltekin. 2014. "An Empirical User Study for Measuring the Influence of Colour Distance and Font Size in Map Reading Using Eye Tracking." *The Cartographic Journal*. Online advance publication. doi:10.1179/1743277414Y.0000000103.

Brychtová, A., and A. Çöltekin. 2015. "Discriminating Classes of Sequential and Qualitative Colour Schemes." *International Journal of Cartography* 1(1): 62–78. doi:10.1080/23729333.2015.1055643.

Brychtová, A., and J. Doležalová. 2015. "Designing Usable Sequential Color Schemes for Geovisualizations." In *Proceedings of the 1st ICA European Symposium on Cartography*, edited by G. Gartner and H. Huang, 31–32. Vienna: Vienna University of Technology. http://eurocarto.org/wp-content/uploads/2015/10/1-7.pdf

Brychtová, A., and A. Vondráková. 2014. "Green versus Red: Eye-Tracking Evaluation of Sequential Colour Schemes." In *SGEM 2014 Informatics, Geoinformatics and Remote*

Sensing Proceedings Volume III, 8. Sofia: STEF92 Technology Ltd. doi:10.5593/SGEM2014/B23/S11.082.

Carter, R., and R. Huertas. 2009. "Ultra-Large Color Difference and Small Subtense." *Color Research & Application* 35 (1): 4–17. doi:10.1002/col.20557.

CIE. 2014. "CIEDE2000 Colour-Difference Formula." In *Colorimetry - Part 6*. Vienna: CIE. http://www.cie.co.at/index.php?i_ca_id=886

Çöltekin, A. 2006. *Foveation for 3D Visualization and Stereo Imaging*. Helsinki: Helsinki University of Technology.

Çöltekin, A. 2009. "Space-Variant Image Coding for Stereoscopic Media." In *2009 Picture Coding Symposium, PCS 2009*, edited by E. J. Delp and T. S. Huang, 1–4. New York: IEEE. doi:10.1109/PCS.2009.5167396.

Cromley, R. G. 1989. "Classed versus Unclassed Choropleth Maps: A Question of How Many Classes." *Cartographica: The International Journal for Geographic Information and Geovisualization* 32 (4): 15–27. doi:10.3138/J610-13NU-5537-0483.

Cybulski, P. 2014. "Rotating Point Symbols on Animated Maps for the Presentation of Quantitative Data." *Kartographische Nachrichten* 4/2014: 198–203.

Dall'Acqua, L., A. Çöltekin, and J. Noetzli. 2013. "A Comparative User Evaluation of Six Alternative Permafrost Visualizations for Reading and Interpreting Temperature Information." Paper presented at GeoViz Hamburg 2013, HafenCity University, Hamburg, Germany, March 6–8. http://www.geomatik-hamburg.de/geoviz/online/dallacqua_Abstract.pdf

Danilova, M. V., and J. D. Mollon. 2006. "The Comparison of Spatially Separated Colours." *Vision Research* 46 (6–7): 823–836. doi:10.1016/j.visres.2005.09.026.

Danilova, M. V., and J. D. Mollon. 2010. "Parafoveal Color Discrimination: A Chromaticity Locus of Enhanced Discrimination." *Journal of Vision* 10 (1): 4.1–9. doi:10.1167/10.1.4.

Dent, B. D., J. S. Torguson, and T. W. Hodler. 2009. *Cartography: Thematic Map Design*. 6th ed. New York: McGraw Hill Education.

Derefeldt, G., T. Swartling, U. Berggrund, and P. Bodrogi. 2004. "Cognitive Color." *Color Research & Application* 29: 7–19. doi:10.1002/col.10209.

Foster, D. H. 2011. "Color Constancy." *Vision Research* 51: 674–700. doi:10.1016/j.visres.2010.09.006.

Garlandini, S., and S. I. Fabrikant. 2009. "Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization." In *COSIT 2009*. Lecture Notes in Computer Science 5756, edited by S. K. Hornsby, 195–211. Berlin: Springer-Verlag. doi:10.1007/978-3-642-03832-7_12.

Gegenfurtner, K. R., M. Bloj, and M. Toscani. 2015. "The Many Colours of 'The Dress.'." *Current Biology* 25: R543-R544. doi:10.1016/j.cub.2015.04.043.

Gilmartin, P. 1981. "Influences of Map Context on Circle Perception." *Annals of the Association of American Geographers* 71 (2): 253–258. doi:10.1111/j.1467-8306.1981.tb01351.x.

Goldberg, J. H., and X. P. Kotval. 1999. "Computer Interface Evaluation Using Eye Movements: Methods and Constructs." *International Journal of Industrial Ergonomics* 24: 631–645. doi:10.1016/S0169-8141(98)00068-7.

Gordon, J., and I. Abramov. 1977. "Color Vision in the Peripheral Retina II Hue and Saturation." *Journal of the Optical Society of America* 67 (2): 202. doi:10.1364/JOSA.67.000202.

Harrower, M., and M. Bloch. 2006. "Mapshaper.Org: A Map Generalization Web Service." *IEEE Computer Graphics and Applications* 26 (4): 22–27. doi:10.1109/MCG.2006.85.

Harrower, M., and C. A. Brewer. 2011. "ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps." In *The Map Reader: Theories of Mapping Practice and Cartographic Representation*, edited by M. Dodge, R. Kitchin, and C. Perkins, 261–268. Hoboken, NJ: Wiley Online Library. doi:10.1002/9780470979587.ch34.

Holmqvist, K., M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van De Weijer. 2011. *Eye Tracking: A Comprehensive Guide to Methods and Measures*. 1 ed. Oxford: Oxford University Press.

ICC. 2014. "Introduction to the ICC Profile Format." *International Color Consortium*. http://www.color.org/iccprofile.xalter

Ishiguro, Y., and J. Rekimoto. 2011. "Peripheral Vision Annotation. In *Proceedings of the 2nd Augmented Human International Conference - AH '11*, 1–5. New York: ACM Press. doi:10.1145/1959826.1959834

Jenny, B., and N. V. Kelso. 2007. "Color Design for the Color Vision Impaired." *Cartographic Perspectives* (58): 61–67. doi:10.14714/CP58.270.

Knapp, L. 1995. "A Task Analysis Approach to the Visualization of Geographic Data." In *Cognitive Aspects of Human Computer Interaction for Geographic Information Systems*, edited by T. L. Nyerges, D. M. Mark, R. Laurini, and M. J. Egenhofer, pp. 355–371. Dordrecht: Springer. doi:10.1007/978-94-011-0103-5_25.

Lafer-Sousa, R., K. L. Hermann, and B. R. Conway. 2015. "Striking Individual Differences in Color Perception Uncovered by 'The Dress' Photograph." *Current Biology* 25: R545-R546. doi:10.1016/j.cub.2015.04.053.

Lenneberg, E. H. 1961. "Color Naming, Color Recognition, Color Discrimination: A Re-Appraisal." *Perceptual and Motor Skills* 12: 375–382. http://www.amsciepub.com/doi/pdf/10.2466/pms.1961.12.3.375.

May, M. 2009. *Sensation and Perception*. New York: Chelsea House.

Meihoefer, H. J. 1969. "The Utility of the Circle as an Effective Cartographic Symbol." *The Canadian Cartographer* 6 (2): 105–117. doi:10.3138/J04Q-1K34-26X1-7244.

Mersey, J. E. 1990. "Colour and Thematic Map Design: The Role of Colour Scheme and Map Complexity in Choropleth Map Communication." *Cartographica* 27 (3): 1–167.

Nagy, A. L., R. R. Sanchez, and T. C. Hughes. 1990. "Visual Search for Color Differences with Foveal and Peripheral Vision." *Journal of the Optical Society of America A* 7 (10): 1995–2001. doi:10.1364/JOSAA.7.001995.

Olson, J. M. 1987. "Color and the Computer in Cartography." In *Color and the Computer*, edited by H. J. Durrett, 205–219. San Diego, CA: Academic Press.

Olson, J. M., and C. A. Brewer. 1997. "An Evaluation of Color Selections to Accommodate Map Users with Color-Vision Impairments." *Annals of the Association of American Geographers* 87 (1): 103–134. doi:10.1111/0004-5608.00043.

Özgen, E. 2004. "Language, Learning, and Color Perception." *Current Directions in Psychological Science* 13 (3): 95–98. doi:10.1111/j.0963-7214.2004.00282.x.

Pele, O., and Werman, M. 2012. "Improving Perceptual Color Difference using Basic Color Terms." *Computer Research Repository* abs/1211.5 (November): 1–14.

Phillips, R. J., and L. Noyes. 1982. "An Investigation of Visual Clutter in the Topographic Base of a Geological Map." *The Cartographic Journal* 19 (2): 122–132. doi:10.1179/caj.1982.19.2.122.

Poole, A., and L. J. Ball. 2005. "Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future." In *Encyclopedia of Human-Computer Interaction*, edited by C. Ghaoui, 211-219. Hershey, PA: IGI Global. doi: 10.4018/978-1-59140-562-7.ch034.

Popelka, S. 2014. "Optimal Eye Fixation Detection Settings For Cartographic Purposes." In *SGEM 2014 Informatics, Geoinformatics and Remote Sensing Proceedings*, Volume I, 705–712. Sofia: STEF92 Technology Ltd. doi:10.5593/SGEM2014/B21/S8.091.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rosenholtz, R., L. Yuanzhen, and L. Nakano. 2007. "Measuring Visual Clutter." *Journal of Vision* 7: 17–17. doi:10.1167/7.2.17.

Roy, M. S., M. J. Podgor, B. Collier, and R. D. Gunkel. 1991. "Color Vision and Age in a Normal North American Population." *Albrecht von Graefes Archiv Fur Klinische Und Experimentelle Ophthalmologie* 229 (2): 139–144. doi:10.1007/BF00170545.

Rubin, J., and D. Chisnell. 2008. *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. 2nd ed. Hoboken, NJ: Wiley.

SensoMotoric Instruments. 2013. *SMI Experiment Center™*. http://www.smivision.com/en/gaze-and-eye-tracking-systems/products/experiment-center-software.html

Sharma, G., W. Wu, and E. N. Dalal. 2005. "The CIEDE2000 Color-Difference Formula: Implementation Notes, Supplementary Test Data, and Mathematical Observations." *Color Research & Application* 30 (1): 21–30. doi:10.1002/col.20070.

Sheppard, S. R. J., A. Shaw, D. Flanders, and S. Burch. 2008. "Can Visualisation Save the World?–Lessons for Landscape Architects from Visualizing Local Climate Change." In *Digital Design in Landscape Architecture*, edited by E Buhmann, M. Pietsch, E. Kretzler, J. Kieferle and P. Paar, 29–31. Heidelberg, Germany: Wichmann Verlag.

Stone, M. 2012. "In Color Perception, Size Matters." *IEEE Computer Graphics and Applications* 32 (2): 8–13. doi:10.1109/MCG.2012.37.

Stone, M., D. A. Szafir, and V. Setlur. 2014. An Engineering Model for Color Difference as a Function of Size. In *Color and Imaging Conference*, 253–258. Society for Imaging Science and Technology. https://research.tableau.com/sites/default/files/2014CIC_48_Stone_v3.pdf

The LimeSurvey project. 2011. *Limesurvey - the Free and Open Source Survey Software Tool*. https://www.limesurvey.org

Vande Velde, L., S. Chatzinotas, M. Larson, J. Löffler, and G. Paliouras. 2009. Interactive 2D - 3D Digital Maps for the Support of Emergency Teams during Rescue Operations. In *12th World Congress on Intelligent Transport Systems 2005*, Vol. 3, 1575–1584. Washington DC: Intelligent Transportation Society of America.

Voßkühler, A. 2013. *OGAMA (Opengazeandmouseanalyzer)*. Berlin: Freie Universität Berlin. Accessed February 4, 2016, http://www.ogama.net/

Whitney, D., and D. M. Levi. 2011. "Visual Crowding: A Fundamental Limit on Conscious Perception and Object Recognition." *Trends in Cognitive Sciences* 15 (4): 160–168. doi:10.1016/j.tics.2011.02.005.

Winkler, A. D., L. Spillmann, J. S. Werner, and M. A. Webster. 2015. "Asymmetries in Blue–Yellow Color Perception and in the Color of 'The Dress'." *Current Biology* 25: R547-R548. doi:10.1016/j.cub.2015.05.004.

Withouck, M., K. A. G. Smet, and P. Hanselaer. 2015. "Brightness Prediction of Different Sized Unrelated Self-Luminous Stimuli." *Optics Express* 23 (10): 13455. doi:10.1364/OE.23.013455.

Wolfe, J. M. 1994. "Guided Search 2.0 A Revised Model of Visual Search." *Psychonomic Bulletin & Review* 1 (2): 202–238. doi:10.3758/BF03200774.

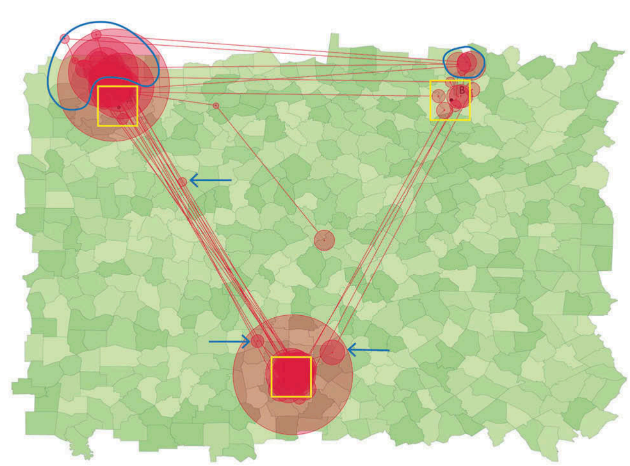
Wong, B. 2011. "Points of View: Color Blindness." *Nature Methods* 8 (6): 441–441. doi:10.1038/nmeth.1618.

Yang, Y., J. Ming, and N. Yu. 2012. "Color Image Quality Assessment Based on CIEDE2000." *Advances in Multimedia* 2012: 1–6. doi:10.1155/2012/273723.

Appendixes

Appendix 1

Blue arrows point to "interfixations" that are mentioned in Section 4.3 and 4.5. We believe that sometimes participants are not able to move their eyes directly to the target, because it may be too far. However, another issue, which contributes to transitions from A↔X, B↔X, or C↔X, is also the accuracy of the eye tracker. In the image below, the yellow rectangles depict AOI A, B, and C; the rest of the picture is X. Blue "bubbles" point to fixations, which must have been intended for the depicted areas, but did not enter the AOI, simply because of the eye tracker precision and/or calibration errors.



Appendix 2

Tentative view of sequential color schemes created with Sequential Color Scheme Generator 1.0 (http://eyetracking.upol.cz/color/). Shades of these schemes are distinguished with constant color difference of $\Delta E_{00} = 10$.

The figure demonstrates that it is not possible to create equal number of classes for all colors across the whole spectrum; that is, we can only do four classes (of $\Delta E_{00} = 10$) from yellow (RGB 255, 255, 0) to white, while blue (starting with RGB 0, 65, 170) gives us seven classes.

RGB codes relate to the sRGB color space. Conversion from RGB to Lab (which is important for further color distance computation) was done with respect to the illuminant D65 and 10°observer.