# Choosing Colors for Map Display Icons Using Models of Visual Search

**Joshua Shive**, Tennessee State University, Nashville, Tennessee, and
**Gregory Francis**, Purdue University, West Lafayette, Indiana

**Objective:** We show how to choose colors for icons on maps to minimize search time using predictions of a model of visual search.

**Background:** The model analyzes digital images of a search target (an icon on a map) and a search display (the map containing the icon) and predicts search time as a function of target-distractor color distinctiveness and target eccentricity.

**Method:** We parameterized the model using data from a visual search task and performed a series of optimization tasks to test the model's ability to choose colors for icons to minimize search time across icons. Map display designs made by this procedure were tested experimentally. In a follow-up experiment, we examined the model's flexibility to assign colors in novel search situations.

**Results:** The model fits human performance, performs well on the optimization tasks, and can choose colors for icons on maps with novel stimuli to minimize search time without requiring additional model parameter fitting.

**Conclusion:** Models of visual search can suggest color choices that produce search time reductions for display icons.

**Application:** Designers should consider constructing visual search models as a low-cost method of evaluating color assignments.

**Keywords:** quantitative modeling, optimization, histogram backprojection, interface design, CIELAB

Address correspondence to Joshua Shive, Tennessee State University, Department of Psychology, 3500 John A. Merritt Blvd., Nashville, TN 37029; e-mail: jshive@tnstate.edu.

*HUMAN FACTORS*
Vol. 55, No. 2, April 2013, pp. 373-396
DOI:10.1177/0018720812459341


## INTRODUCTION

Color coding reduces search time for items on visual displays (Phillips, 1979; Van Laar & Deshe, 2002; Williams, 1967; Yeh & Wickens, 2001). Color codes exist for certain types of items (Dent, 1990; Harvey, 2008), such as on maps, where icons representing the locations of emergency services are usually red. A person searching for a first-aid station on a map that employs this color code can restrict his or her search to red icons only, thereby reducing search time. For other icons, color code conventions may be nonexistent. The goal of this study is to determine what colors a designer should assign these icons to minimize search time.

Our current project shows how to use a model of visual search to choose colors for icons on maps. We first review color-coding guidelines and describe the factors that determine visual search performance. Then, we describe how to construct a quantitative model that incorporates those factors and fits data from an experiment (Experiment 1) in which participants search for colored icons on maps. We show that the model provides a strong fit to the Experiment 1 data, indicating that the model can predict search times for specific colored icons.

Next, we describe a procedure for choosing the optimal color assignment for a set of icons on a map. The optimization algorithm facilitates consideration of a very large number of possible color assignments to find the model-predicted optimal assignment. We test the parameterized model's ability to choose colors for icons on maps to optimize for a number of characteristics. First, we show how to choose color assignments to minimize average search time across icons when some icons are considered more important than others. We test the predicted optimal designs with human observers (Experiment 2) and conclude that this approach can facilitate search for important icons at the expense of search for other icons.

Next, we explore the model's flexibility by having it choose optimal color assignments for a different set of icons, colors, and background maps than the ones used to set its parameters. Again, we compare the model's predictions with human search performance (Experiment 3). We show that this approach can facilitate search for important icons even when the model has not previously been exposed to a set of icons or colors. Then, we show how to use the model to choose color assignments that minimize other characteristics, such as the predicted standard deviation of search times across icons (i.e., making every icon as easy to find as every other icon). The results indicate that the optimal color assignment for a set of map icons depends on the characteristic being optimized. Finally, we compare our approach with alternative methods of choosing colors.

## Color Coding on Maps

Maps are visual representations of spatial data. Kraak and Ormeling (1996) identify two broad types of maps. Topographic maps show the features (e.g., rivers, roads, buildings, relief) of a section of the earth's surface. Thematic maps present the spatial distribution of a particular phenomenon, such as the most common political affiliation for voters in each county of a state. Maps aid a variety of tasks. Most maps (local maps, terrain maps, aviation sectional charts, etc.) are used for orientation and navigation. Organizations also use maps for physical planning, management, and monitoring purposes. For example, a city will use different maps to anticipate future land use, show the zoning of different areas of town, and keep track of which roadways are most in need of maintenance.

Color can be used to convey information about the relationships between categories of data shown on a map and can draw attention to important elements in a visual display. Cartographers suggest considering the measurement scale of the data when choosing colors for data categories (Brewer, 2005). On a thematic map showing geographical information about counties in a state, color can be used to show qualitative differences (areas of plain, desert, forest, etc.) or sequential differences (average yearly rainfall). Brewer cautions against choosing color schemes that imply an order of categories when order does not exist. For example, a map showing geographical information might use a color palette of different hues to represent different types of terrain. However, it would be inappropriate to use a "hot-to-cold" color scheme to represent the different categories, since that might imply that terrain types are ordered in a logical fashion.

Map design guidelines suggest varying color dimensions, such as hue, saturation, and value (lightness) to create colors that represent categories. One can create qualitative color schemes by choosing distinct hues for different categories. Sequential color schemes involve changes in lightness for the creation of categories. Harrower and Brewer (2003) created an online tool, ColorBrewer.org, that allows designers to compare color schemes for different varieties of thematic maps. With ColorBrewer.org, users specify the number of data categories and the measurement scale of the data, and the tool generates a number of possible color schemes for representing the data. Users can preview potential color schemes as they would appear on a map and download the proposed color schemes in a number of color specification systems.

A related line of color-coding research (Carter & Carter, 1982; Campadelli, Posenato, & Schettini, 1999) identifies sets of maximally discriminable colors for qualitative categories. In perceptually based color spaces, such as CIELAB or CIELUV, the distance between two colors in three-dimensional space corresponds to the discriminability of the two colors (Fairchild, 1998). R. C. Carter and E. C. Carter (1982) describe an algorithm for selecting a set of high-contrast colors that involves maximizing the minimum distance between any two colors in a perceptually based color space.

Color can be used to segment portions of a visual display to facilitate search (Grossberg, Mingolla, & Ross, 1994; Yeh & Wickens, 2001). Van Laar and Deshe (2002) showed observers a control panel display in which items that belonged to the same category were presented in the same hue as the background on which they sat. Background hues were less saturated than the hues in which items were presented, creating the appearance of visual layers. Observers were asked to search for information presented in different areas of the display. Searches were faster with the layered display than with either a monochrome display or a display in which items were presented with the use of maximally distinct colors.

Van Laar and Deshe's (2002) results demonstrate that although it is important to choose display colors that are distinctive from one another, the designer who wants to minimize visual search for items presented on displays must also consider the factors that influence search time. In this article, we choose optimal colors for maps using a model that predicts how long it will take an observer to search for a particular target on a display. The model includes factors known to influence search time, such as the distinctiveness of the target's features and the location of the target on the display relative to the observer's initial fixation. We review these factors in the next section.

## Modeling Visual Search for Maps

In a typical visual search task, a participant examines a display containing a number of items, looking for a target that may or may not be present. If the target is present, the participant makes one response, such as a key press. If the target is absent, the participant presses another key. Typically, both search time and accuracy are measured.

Psychologists studying visual search have described what types of displays produce fast or slow searches (Treisman & Gelade, 1980), examined the relationship between search time and eye movements (Zelinsky, 1997), and proposed a number of models for predicting visual search behavior in complex scenes. Some of these models predict patterns of eye movements to salient regions of a display (Itti & Koch, 2000) or to locations on a display that may contain a target item (Hwang, Higgins, & Pomplun, 2009; Navalpakkam & Itti, 2007; Wolfe, 2007; Zelinsky, 2008). Other models predict visual search time on the basis of the amount of clutter a visual display contains (Rosenholtz, Li, & Nakano, 2007; Bravo & Farid, 2008). Still other models involve consideration of optimal visual search strategies observers use to minimize search (Najemnik & Geisler, 2005).

Although these models predict search performance in a variety of situations, none of them is suited for the current project. Some are purely bottom-up (Itti & Koch, 2000; Rosenholtz et al., 2007). Others predict patterns of eye movements rather than search times (Hwang et al., 2009; Najemnik & Geisler, 2005; Zelinsky, 2008). Still others are not described in sufficient detail to implement with real images (Navalpakkam & Itti, 2007; Wolfe, 2007). For these reasons, we developed our own visual search model.

In the current study, we model visual search time in a task whereby an observer searches for a colored icon representing a location in a park (i.e., the target) on a display containing other icons (i.e., distractors) and background features representing roads and rivers. We hypothesize that search time will be influenced by two primary factors: target-distractor color discriminability (*C* in the model) and target eccentricity (*E* in the model). Other factors are known to affect search performance. However, with too many model factors, it may be difficult to interpret the optimization's color choices. As a result, we selected a single color factor that seemed reasonable. In addition, prior research in our lab (Shive & Francis, 2008; Francis, Bias, & Shive, 2010) had revealed a strong eccentricity effect in the visual search tasks we use to study search time. Thus, we included target eccentricity in the model. Before describing how to construct the model, we review research examining each of these factors.

When a search target has a feature not shared by any other items, search is fast, regardless of the number of distracting items (Yantis, 1993). When the target must be distinguished from distracting items on the basis of a combination of features, search time increases as the number of distracting items that share the target's features increases (Treisman & Gelade, 1980).

Even if distracting items on a display do not share a target's exact features, the similarity between the target and distractors affects search time (Desimone & Duncan, 1995; Duncan & Humphreys, 1989, 1990). When the target has features that are highly discriminable from those of distracting items, search time is reduced. For example, search for a red target among green distractors is easier than search for a red target among pink distractors. Similarly,

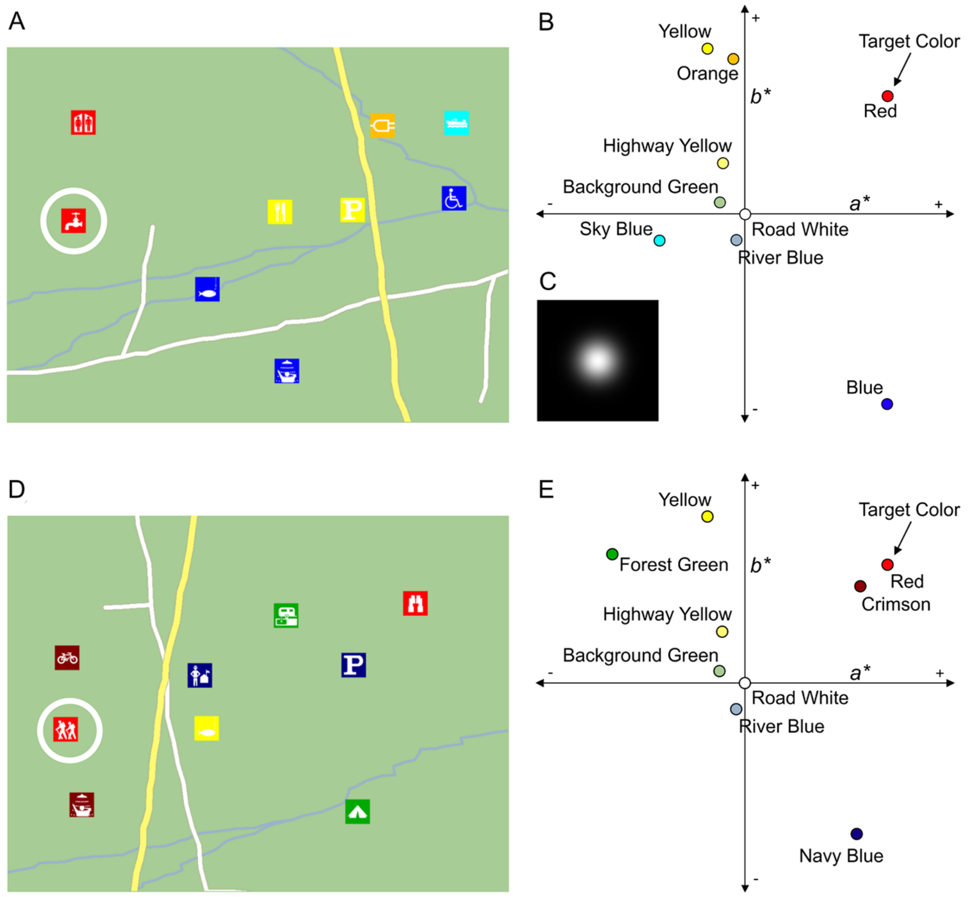


*Figure 1.* (A) Example trial from Experiment 1. The target is a red icon (indicated by a white circle) and one red distractor is present. (B) The ($a^*$, $b^*$) color space distribution for the trial. The red pixels in the target icon and distractor icon are located in the upper-right-hand corner of the distribution. (C) The Gaussian filter used to blur the display histogram (see appendix for details). (D) Example trial from Experiment 1. The target is a red icon. One red distractor and two crimson distractors are also present. (E) The ($a^*$, $b^*$) color space distribution for the trial shown in (D). The red target, the red distractor, and the crimson distractors are all located in the upper-right-hand corner of the distribution.

search for a *T* among rotated *T*s becomes more difficult the closer the distracting *T*s are rotated to vertical. Thus in Figure 1, the circled red icon on the upper map should be found faster than the circled red icon on the lower map, because the upper map contains only one red distracting icon, whereas the lower map contains one red distracting icon and two crimson icons.

The effects of target-distractor discriminability extend to features of a display's background (Rosenholtz, 2001). However, models that fit data from visual search experiments tend to ignore the background in representing the search scene. For example, Guided Search (Wolfe, 2007) represents each search display as a set of discrete items with a limited list of features. Our maps are images containing many background features that are not easily represented in this manner, so we developed a measure of target-distractor color discriminability to analyze all pixels in a search display and to compare the features at each pixel with the pixels in the search target's image. This approach is similar to the Feature Congestion measure

*Figure 2.* Icons used in Experiment 1.

described by Rosenholtz et al. (2007), which discards the concept of set size (number of display items) in favor of a general measure of display clutter. (It should be noted that our measure generates search time predictions for specific targets, whereas the Feature Congestion measure does not.) We hypothesized that search time would decrease as target-distractor color discriminability increased.

Icons closer to the center of a search display receive more eye fixations (Zelinsky, 1997) and are found more quickly than icons located farther from the center of the display (Shive & Francis, 2008; Williams, Reingold, Moscovitch, & Behrmann, 1997; Scialfa & Joffe, 1998). The current model calculates target eccentricity as the distance from the center of the display to the center of a target icon. We predicted that search time would increase as eccentricity increased. Thus, on the upper map shown in Figure 1, the yellow icon located just to the right of the center of the display should be found faster than the circled icon.

We propose that the time required to find an icon on a map can be modeled as

$$P_i = \alpha + B_1 C_i + B_2 E_i, \tag{1}$$

where $P_i$ is the predicted search time for trial $i$, $\alpha$ is the intercept, $B_1$ is the weight of target-distractor color distinctiveness, $C_i$ is the target-distractor color distinctiveness for trial $i$, $B_2$ is the weight of target eccentricity, and $E_i$ is the target's eccentricity for trial $i$. The appendix defines target-distractor color distinctiveness quantitatively.

## EXPERIMENT 1: PARAMETERIZING THE MODEL

We conducted a visual search experiment to gather data that we used to set the parameters of the model. To produce the best model fit to the data, we designed the stimuli and procedure of Experiment 1 to gather as many observations of search time as possible using stimuli that varied as much as possible.

### Method

*Stimuli*. We collected a set of icons representing locations of interest (restrooms, swimming areas, parking lots, etc.) in an imaginary state park. Figure 2 shows examples of the icons used in Experiment 1. We selected multiple icons that represented each location of interest. As Figure 2 shows, each icon consisted of a white and black pictogram centered on a black square. It should be noted that even though we used square icons with superimposed symbols, the model we describe can predict search time for icons of any shape. (Throughout the manuscript, we use the term *icon* to refer to the combined area containing all of the pixels in the pictogram and the surrounding region. We use the term *background* to refer only to the map display on which the icons were placed.)

Experiment 1 required us to obtain a large number of maps. Rather than scanning maps from books or taking screenshots of maps from the Internet, we decided to create the maps from scratch using a computer program. We selected the background features of the maps on the basis of the color scheme Google Maps uses for state and national parks. We chose this color scheme because it is relatively simple (i.e., easy to incorporate in a computer program designed to generate hundreds of maps). In addition, the Google Maps scheme follows several common color code conventions (e.g., water is blue, forests are green, and different colors are used for interstates, highways, and streets). However, different maps use different color schemes. It is unknown how the model's fit will change if the color scheme is changed.

We created the maps using the drawing tools in the Psychophysics Toolbox in MATLAB (Brainard, 1997; Pelli, 1997). Each map contained

**TABLE 1:** Colors of Icons in Experiment 1 and Experiment 3

| Experiment 1: Park Maps | | Experiment 3: City Maps | |
|---|---|---|---|
| Color | RGB Value | Color | RGB Value |
| Crimson | 128,0,0 | Blue-gray | 0,128,192 |
| Red | 255,0,0 | Dark orange | 255,128,0 |
| Orange | 255,191,0 | Grape | 128,128,192 |
| Yellow | 255,255,0 | Fuchsia | 255,0,128 |
| Lime green | 79.688,255,0 | Gold | 128,128,0 |
| Forest green | 0,168,0 | Aquamarine | 62,181,130 |
| Sky blue | 0,255,255 | Brown | 130,65,0 |
| Navy blue | 0,0,128 | Red-brown | 128,64,64 |
| Blue | 0,0,255 | Mint green | 0,198,0 |
| Purple | 223.13,0,255 | Salmon | 255,128,128 |

nine colored icons as well as background features representing roads, rivers, and green space. Each map display measured 768 pixels × 1,024 pixels. A total of 10 possible colors were used for the display icons. Colors were selected from the default MATLAB color map. The first two columns of Table 1 list the colors and their RGB values. Displays were presented on a monitor calibrated to display standard colors. Monitor calibration was performed with a Pantone Eye-One calibrator.

The icons were placed on a green background (RGB: 167,204,149) containing roads (RGB: 255,255,255) and rivers (RGB: 153,179,204). To create each map, we first selected the colors of icons randomly from the colors listed in the first two columns of Table 1. No attempt was made to control the number of icons with identical colors or the number of times the target was assigned a particular color.

In state parks, the locations of parking lots, boat docks, and other areas are determined by geography. As a result, the icons that represent those locations rarely appear in straight lines when viewed on a map. The icons' positions on the maps were selected from positions on a 4 × 6 matrix of locations spread over the display to ensure that no icons overlapped. Icon positions were jittered randomly by a value between −50 and +50 pixels both vertically and horizontally so the icons were unlikely to be placed in straight lines with one another. We chose one icon at random from the display as the target for that display.

We asked participants to sit at a comfortable distance from the monitor. This distance varied slightly across individuals. At a distance of approximately 15 inches, the displays occupied approximately 36° × 48° of visual angle. Each icon measured 50 pixels × 50 pixels and occupied approximately 2.38° × 2.38° of visual angle.

We created 450 unique trials and divided them into three groups of 150 trials. A total of 15 observers viewed the first group of trials, 15 observers viewed the second group of trials, and 15 observers viewed the third group.

*Participants*. Observers were from the Purdue University participant pool and received course credit for their participation. Each participant reported normal or corrected-to-normal vision and normal color vision.

*Procedure*. The 150 map displays were randomly ordered. Figure 3 shows the procedure for a single trial. An observer was first shown the icon that was the search target for that trial. The icon appeared in the center of the screen on a white background, as it might be if presented on a map legend. The icon was displayed in the same color and at the same size as it would appear on the map. The target icon appeared for 1,500 ms and was then replaced by the map display, which filled the screen. We instructed observers to find the target as quickly as possible and to press the space bar when they found it.

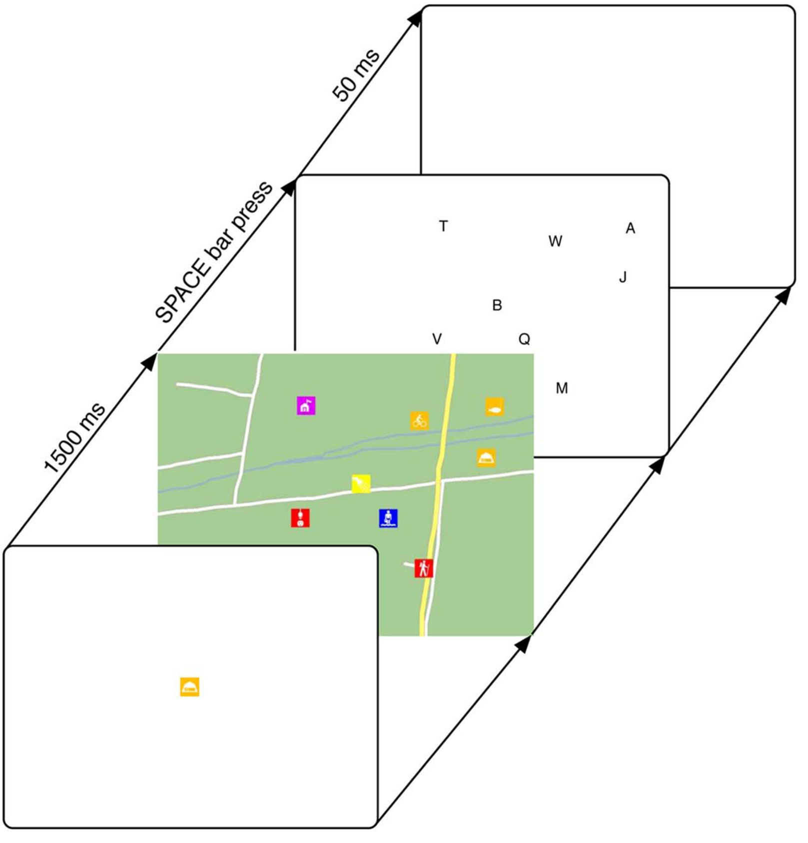


*Figure 3.* Experimental procedure for a single trial from Experiment 1. A participant sees a colored search target and then searches a map display until pressing the space bar to indicate that the target has been found. Next, an array of letters, each one located where a display icon had been located on the map, appears for 50 ms. The participant then presses the key corresponding to the letter that appeared in the target icon's location.

On half of the trials in traditional visual search experiments (e.g., Treisman & Gelade, 1980), the search target is not present on the display. However, since we are evaluating a method for choosing colors to minimize search time for icons on maps, target-absent trials would provide no information that we could use to set the parameters of our search model. As a result, we used a "no-cheat" paradigm whereby the target was present on every trial (Navalpakkam & Itti, 2006, 2007). After the observer pressed the space bar to report that the target for a trial had been found, an array of randomly selected letters was displayed for 50 ms. Each letter in the array was located in the spot where a display icon had appeared on the map for that trial. After 50 ms, a blank background replaced the array. Observers were given an untimed period to press the key corresponding to the letter that had appeared at the target icon's location. This procedure ensured that the observer's attention was directed at the target icon's location when the observer reported finding the target. After responding, the observer received feedback about whether the response was correct or incorrect, and search time was calculated. The experiment took approximately 20 min to complete.

## Results

We calculated average search times across participants for each of the 150 trials for each group. Trials in which observers made incorrect letter responses (1.8%) were discarded from further analysis. Combining across observers, we obtained an estimate of average search time for a target for each of the 450 unique map designs. We used these estimates of search time to fit the parameters of the visual search model.

## Model Construction

To construct a model of visual search that fits search times for the Experiment 1 trials, we calculated the target-distractor color distinctiveness score for each trial using an algorithm called *histogram backprojection*. Details of the algorithm can be found in the appendix. The distinctiveness score was then combined with the eccentricity score for the target. These scores were computed for each of the 450 trials.

Figure 1 shows two trials from Experiment 1 and illustrates how the model's measure of color distinctiveness predicts the difference in search time between the two trials. Figure 1A shows a map on which the target is a red icon (indicated by a white circle). The display also contains one red distractor icon. The average experimental search time for this trial in Experiment 1 was 638.5 ms. Figure 1B shows the distribution of the colors used on the map in the CIELAB color space. The $x$-axis represents the $a*$ dimension of the color space; the $y$-axis represents the $b*$ dimension. For this map, the target's color is the only color in the upper-right-hand quadrant of the color space distribution. Target-distractor color distinctiveness ($C_i$) for this trial is 1060.78. Eccentricity, $E_i$ is 376.04. Model-fitted search time for this trial is 895 ms.

Figure 1D shows a map from Experiment 1 on which the target is a red icon and the display contains a red distractor, as before. However, this display also contains two crimson icons. Average search time for this trial in Experiment 1 was 856.3 ms, about 220 ms longer than the search time for the trial in Figure 1A. The color space distribution for the display, shown in Figure 1E, reveals how the measure of target-distractor color distinctiveness explains the increased search time on this trial. The presence of a color similar to the target color reduces the color distinctiveness score for the red target (689.41 on the current map, compared with 1060.78 on the map in Figure 1A). Eccentricity is 395.56, compared with 376.04 on the previous map. The decrease in color distinctiveness produces a longer model-fitted search time (930.3 ms, approximately 35 ms longer than model-fitted search time for the trial shown in Figure 1A).

**TABLE 2:** Results of Linear Regression With Data From Experiment 1

| Variable | *B* | *SE B* | β |
|---|---|---|---|
| Constant | 871.23 | 20.92 | |
| Color | –.18 | .01 | –.575 |
| Eccentricity | .53 | .05 | .35* |

*Note.* $R^2$ = .47.
*$p < .0001$.

We performed linear regression using average search time as the dependent variable and target-distractor color distinctiveness data and eccentricity data obtained from the algorithm as factors in the regression equation. Table 2 shows the results of the regression analysis. The model fit the experimental search times with $R^2 = .467$. Linear regression revealed significant effects of target-distractor color distinctiveness ($B_1 = -.18$, $t = -16.64$, $p < .0001$) and eccentricity ($B_2 = .53$, $t = 10.13$, $p < .0001$). Color distinctiveness ($\beta_1 = -.575$) had a stronger effect than eccentricity ($\beta_2 = .350$). Now we show how to use the parameterized model to choose icon colors to optimize for different characteristics.

## Construction of Map Designs Based on Search Frequency

It is possible to consider the problem of choosing colors for visual displays as an optimization problem (Liu, Francis, & Salvendy, 2002; Shive & Francis, 2008; Francis et al., 2010). Given a model that predicts visual search time and a potential palette of colors, this approach can reduce average search time for targets on simple and complex displays.

To create designs that would lead to short average predicted search times, we constructed an optimization algorithm in MATLAB. We gave the algorithm the first 150 maps (out of 450 total maps) from the trials of Experiment 1,

**TABLE 3:** Example Search Frequency Distribution

| Icon | Search Frequency |
|---|---|
| Picnic | 8 |
| Overlook | 7 |
| Water | 6 |
| Cycling | 5 |
| Beach | 4 |
| Electricity | 3 |
| Restroom | 2 |
| Camping | 1 |
| Dining | 0 |

minus the colors assigned to the icons. For each map, we gave the algorithm a list of frequencies that indicated, for each icon on the map, the number of times, out of 36 hypothetical trials, each icon would be the target. The search frequency list for each map was generated randomly so that the search frequencies of individual icons differed across maps. (Note that in Experiment 1, there was no such search frequency list and no color assignment based on search frequency. Instead, the target icon for each map in Experiment 1 was chosen randomly and the colors of icons on the Experiment 1 maps were chosen randomly from Table 1.)

Table 3 shows an example search frequency distribution for a map. The left column lists the icons on the map. We used the same nine icons on all the optimized maps (the icons shown in Figure 2). We chose the order of the icons in the search frequency distribution randomly for each map. The right column of Table 3 shows each icon's search frequency. The optimization used the search frequencies to weight the contribution of the predicted search time for each icon when calculating the predicted average search time across icons. Icons with a higher search frequency contribute more to the predicted average search time than do icons with lower search frequency. In this example, the picnic icon has the highest search frequency and contributes eight search times to the average. The dining icon has the lowest search frequency and contributes zero search times to the average.

With nine icon positions and 10 colors, the number of ways to assign colors to icons is $10^9$, or 1 billion. With so many possible combinations, it is not possible to examine predicted search time for all possible combinations and pick the combination that predicts the shortest average search time. Instead, we devised a gradient-descent algorithm for choosing optimal color assignments that evaluates a subset of possible color combinations.

The algorithm chooses a random assignment of colors to icons and predicts the search time for that color assignment using the search model described in the previous section. Next, the algorithm identifies all possible ways of changing the color of one icon and calculates the change in predicted search time that each color change produces. The algorithm selects the color change that produces the greatest decrease in predicted search time. This procedure continues until no possible color changes reduce predicted search time.

The gradient-descent algorithm ran 1,000 times for each map display using different random initial color assignments. Then, the algorithm chose the color combination that led to the shortest predicted mean search time as the predicted optimal design for that display. When more than one color assignment led to the shortest mean search time, the algorithm chose one of the optimal assignments at random. The optimization followed this procedure for each of the 150 maps.

The results of the optimization were similar across the 150 displays. Since the optimization could pick from 10 colors, the optimization assigned a unique color to each of the nine icons. This result makes sense, since the search model predicts longer search times when more than one icon shares the same color. However, although every icon was given a different color, sky blue and purple were assigned most often to high-frequency icons, because these colors are not close to any other colors in color space.

Forest green and lime green were most often assigned to low-frequency icons. One might expect that this assignment is because of the two colors' perceptual similarity to the background green. In fact, these colors are much more saturated than the background green and so are not especially close to the background green in color space. However, the two colors are close to each other, thereby penalizing color assignments that assign the colors to high-frequency search icons. The optimization's assignment of these colors is a result of their similarity to each other and should not be taken as a general guideline that low-frequency icons be green.

Although the optimization's results minimized predicted search time, the difference in predicted search time between the high-frequency and low-frequency icons across the displays was small (25 ms), because the optimization assigned each icon a different color. A power analysis of this effect with $\alpha = .05$, $\beta = .2$, and the standard deviation of search times from Experiment 1 ($\sigma = 337.76$ ms) reveals that data from 862 participants would need to be collected to show whether this difference exists in an experimental study. Since this is not feasible, we modified the optimization procedure to produce color assignments that led to larger differences in predicted search time by restricting the number of colors that optimized displays could contain.

*Optimization with four colors*. We devised a variation on our gradient-descent algorithm whereby we assessed a penalty for color assignments that used more than four colors. We added the penalty to the predicted average search time for a color assignment using the following decision rule:

$$ST_{predicted} = \begin{cases} ST_{predicted} + p(n_{colors} - 4) & \text{if } n_{colors} > 4; \\ ST_{predicted} & \text{otherwise.} \end{cases} \quad (2)$$

where $ST_{predicted}$ is the predicted average search time across icons on a map for a color assignment, $n_{colors}$ is the number of different colors assigned to display icons, and $p$ is the penalty for using more than four colors. As the decision rule shows, a penalty is added to the predicted average search time of color assignments that use more than four colors. The $p(n_{colors} - 4)$ term weights the value of the penalty on the basis of how many more colors than four the color assignment uses. This term allows the optimization algorithm to deal with initial color assignments that use more than four colors. For example, if the initial color assignment uses seven colors, that assignment will receive a penalty. In addition, all potential ways of changing the color of one icon will be penalized. However, color assignments that use only six colors will receive a smaller penalty than will color assignments that use seven colors, so the algorithm will choose one of those assignments. At each step in the optimization, color changes that reduce the penalty will be favored until the optimization arrives at a four-color assignment that is not penalized.

Our choice to restrict the optimized maps to four icon colors is not intended as a statement about the number of colors a map should contain. We chose the four-color limit because it forces the optimization to differentiate between suboptimal and optimal color assignments, rather than just assigning each icon a unique color, which should produce larger differences in predicted search time between high-frequency and low-frequency icons. We could have selected a higher color limit. However, increasing the number of allowed icon colors may reduce predicted search time differences, requiring data from more participants to validate the optimal color assignments.

We selected the value of the penalty ($p = 69$ ms) to balance between two considerations. A value that is too low will not add sufficient cost to color assignments with more than four colors. On the other hand, a value that is too high may cause the optimization algorithm to converge too quickly, potentially causing it to get stuck in local minima. For example, imagine that the optimization produces an assignment that only uses four colors but is not the best assignment for this map, since two icons share a color that is similar to the display's background. Changing the colors of these two icons would require two optimization steps. The first step would change the color of one of the icons to a better color. This change would reduce predicted search time for the display but produce a penalized five-color assignment. The second step would change the color of the second icon to the better color shared by the first icon, removing the penalty for using five colors. If the penalty for color assignments with more than four colors is too high, such a series of changes might be discouraged.

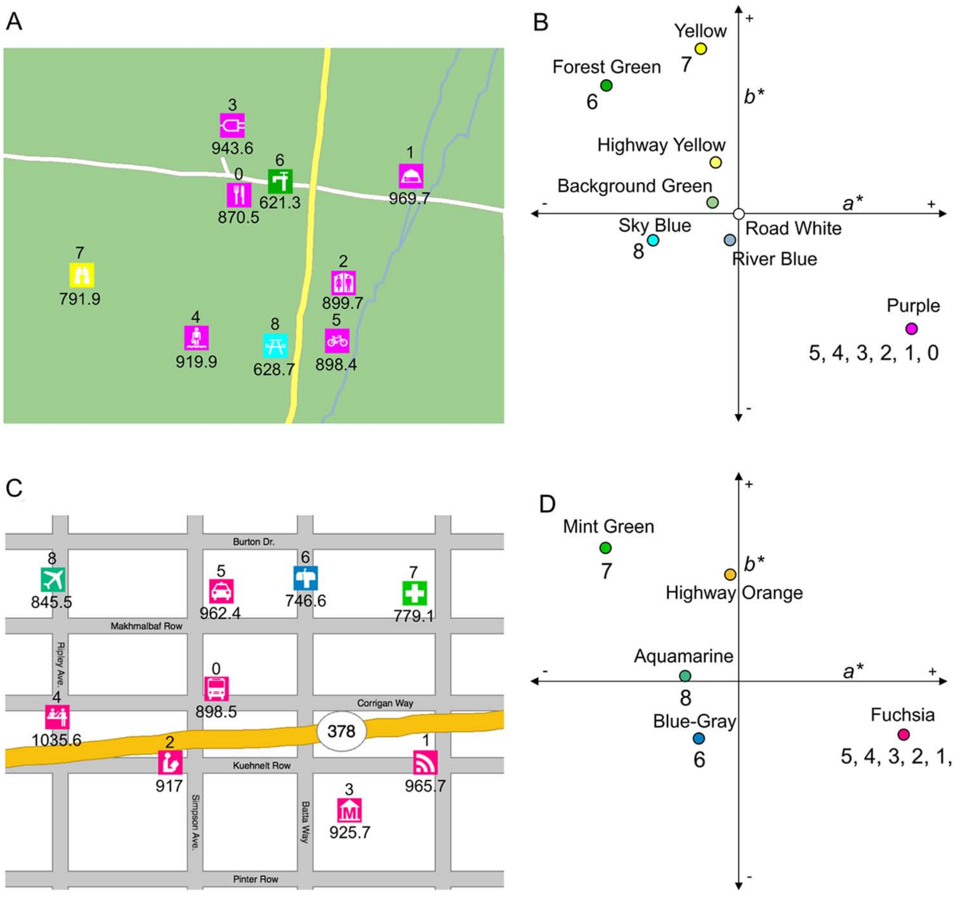


*Figure 4.* (A) Example optimized map. The number above each icon indicates the icon's search frequency. The number below the icon indicates predicted search time (in milliseconds) for that icon. The three icons with the highest search frequencies were assigned unique colors. The rest of the icons were assigned the same color. (B) The ($a^*$, $b^*$) color space distribution for the display in (A). (C) Example four-color optimized city map. The number above each icon indicates the icon's search frequency. The number below the icon indicates predicted search time (in milliseconds) for that icon. The three icons with the highest search frequencies were assigned unique colors. The rest of the icons were assigned the same color. (D) The ($a^*$, $b^*$) color space distribution for the display in (C).

We performed the optimization as before. For each map display, the gradient-descent algorithm considered 1,000 initial color assignments. Each of the 150 map displays from Experiment 1 was assigned the four-color assignment that predicted the shortest average search time across display icons.

*Four-color optimization results*. Figure 4A shows one of the optimized maps. The value above each icon indicates the search frequency for that icon. Icons with higher frequencies contribute more to the weighted average of predicted search times. The value below the icon indicates the predicted search time (in milliseconds) for that icon. The figure illustrates the optimization strategy that produces the minimum predicted search time for each display. The optimization assigns unique colors to the three highest-frequency icons and assigns the same color to all the other icons. High-frequency icons contribute more to the average predicted search time. Search will be fastest for those icons when they have unique colors. The

**TABLE 4:** Search Times for Optimized Park Maps and Optimized City Maps (in milliseconds)

| | | Optimized Maps | | |
|---|---|---|---|---|
| | Original Map | Low Frequency | Medium Frequency | High Frequency |
| Experiment 1 | | | | |
| Experimental | 810.46 | | | |
| Predicted | 810.90 | | | |
| Experiment 2: Park maps | | | | |
| Experimental | | 1196.04 | 1203.96 | 752.23 |
| Predicted | | 932.45 | 944.99 | 703.73 |
| Experiment 3: City maps | | | | |
| Experimental | | 1120.77 | 1193.12 | 845.67 |
| Predicted | | 945.73 | 959.15 | 733.412 |

optimization assigns the rest of the icons the same color to avoid the penalty for using more than four colors but picks a color far away from the other colors in color space. Figure 4B illustrates this principle. The six icons with the lowest search frequencies are assigned the purple color in the lower right quadrant of the color space distribution. This color is far from any of the other display colors in CIELAB color space. The models suggest that this type of color assignment is a good way to minimize average search time. We tested this prediction in Experiment 2.

## EXPERIMENT 2

We conducted an experiment to test the effect of the optimized color assignments on search time. We divided the 150 optimized maps into three groups. The target for the first 50 maps was the icon with the highest search frequency relative to other icons in the optimization (the icon searched for eight times out of 36 hypothetical trials in the optimization). The target for the second 50 maps was an icon with medium search frequency relative to other icons (the icon searched for four times in the optimization). The target for the last 50 maps was the icon with the lowest search frequency relative to other icons (the icon searched for zero times in the optimization). (Even though the icons had different search frequencies in the optimization, in Experiment 2, each map and the target for that map were presented only once. Thus, in the experiment, the true search frequency for each target is one.)

Table 4 shows the predicted search times for the low-frequency (zero searches), medium-frequency (four searches), and high-frequency (eight searches) search trials for the optimized maps. As the table shows, the shortest search times were predicted for the high-frequency search icons, which have a color shared by no other icon on the display. Longer search times were predicted for medium-frequency and low-frequency icons, which share the same color.

The model predicted that average search times for medium-frequency and low-frequency search icons would not differ significantly. Table 4 shows that the predicted average search time for the medium-frequency icons was slightly longer (12 ms) than the predicted average search time for the low-frequency icons. At first, this difference may seem to contradict our prediction that higher-frequency icons should be found faster than lower-frequency icons. However, this predicted difference is not statistically significant and is caused by the locations of icons on the Experiment 1 maps used in the optimization. We did not change the placement of icons on these maps before running the optimization for each map. By chance, the icons from the Experiment 1 maps assigned to be medium-frequency icons in the optimization were 23 pixels farther on average from the center of the display than were the low-frequency

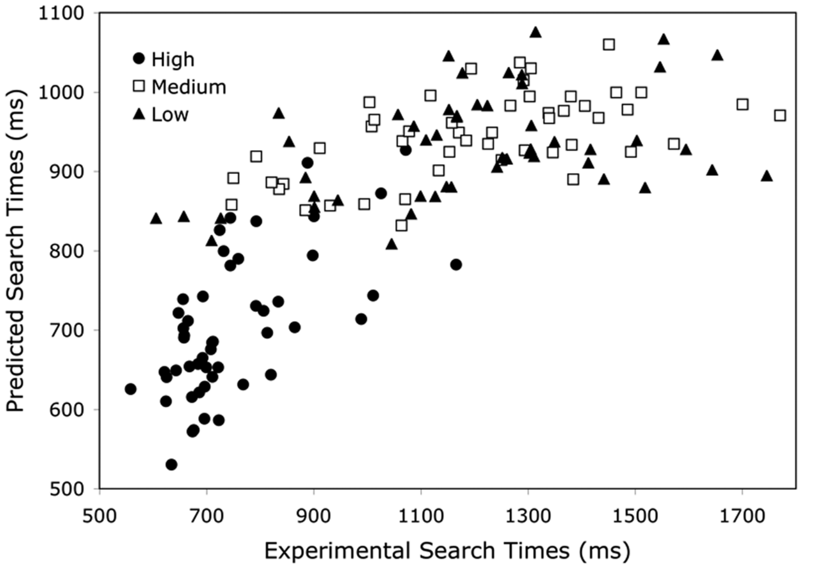


*Figure 5.* Comparison of predicted and experimental search times from Experiment 2 ($R^2$ = .59).

icons, increasing the eccentricity values for the medium-frequency icons. This apparent difference in eccentricities is not statistically significant, and we predict that there should be no empirical difference between search times for medium-frequency icons and low-frequency icons, since the icons share the same color.

## Method

*Procedure*. The experiment followed the same procedure as Experiment 1. Observers saw 5 practice trials followed by the experimental trials. The 150 trials were presented in a random order. Participants were 14 undergraduates from the Purdue University participant pool who took part in the experiment for course credit.

*Results*. We calculated average search times across participants for each of the 150 search trials. Participants made incorrect responses on 2.14% of the trials.

An analysis of variance revealed a significant effect of search frequency, $F(2, 2052)$ = 225.93, $p < .001$. Post hoc analyses with use of the Scheffé criterion for significance indicated that high-frequency search targets were found faster than medium-frequency or low-frequency search targets ($p < .05$ for both comparisons). However, the difference between search times for medium-frequency and low-frequency search targets was not significant.

The model strongly predicted the experimental results ($R^2 = .59$, $p < .001$). Figure 5 shows a scatterplot of the predicted and empirical data. As the figure shows, high-frequency search targets were found faster than medium-frequency targets and low-frequency targets, in line with the model's predictions. In addition, the model predicts that not all high-frequency targets will be found faster than medium-frequency and low-frequency targets. For example, search for a high-frequency search target located at the edge of a display may take longer than search for a medium-frequency target located near the center of the display.

The first two rows of Table 4 compare predicted and experimental search times for the low-, medium-, and high-frequency trials from Experiment 2 with the average fitted and experimental results from Experiment 1. Search times for high-frequency trials in Experiment 2 ($M$ = 752.23 ms) were shorter than the mean search time for Experiment 1 (810.46 ms). This finding illustrates one effect of the optimization. Searches for high-frequency icons in Experiment 2 took less time (58 ms on average) than the average search time for targets in the Experiment 1 trials had taken.

However, the table also shows that searches for high-frequency icons in Experiment 2 were facilitated at the cost of searches for low- and medium-frequency icons. Mean search times for medium-frequency and low-frequency trials (1203.96 ms and 1196.04 ms) were longer than the mean search time for targets on Experiment 1 (810.46 ms). In addition, Table 4 shows that the model correctly predicted that the average search time for low-frequency icons would be slightly faster than the average search time for medium-frequency icons because of the random effects of icon placement on the maps.

The effect of the optimization on searches for high-frequency icons appears small when we consider the reduction in search time from Experiment 1 to Experiment 2 (58 ms, a 7% decrease). Considering more color combinations in the optimization algorithm might produce assignments with faster search times for these icons. However, it is not clear how much more benefit would be gained from this approach. Nevertheless, the effects of optimization may not be as small as they initially appear. Although 58 ms might seem like a small benefit for a single search, the savings across multiple searches by multiple observers may be considerable. In addition, a portion of each measured search time is devoted to nonsearch components, such as the time needed to program and execute a key press. Thus, the proportional benefit to the perceptual and cognitive components of search is likely greater than 7%.

The empirical differences in the data are consistent with the predicted effects of optimization with the model. In general, the optimization approach biases the display to promote the fastest search for targets that are searched most often. The four-color requirement favors color assignments that give targets that are searched most often unique colors. The model accurately predicts what types of displays will generate this pattern of search.

## Assigning Colors to Novel Maps to Minimize Search Time

To test the model's ability to use one set of regression weights to predict performance in a new situation, we gathered empirical data from a new set of targets on optimized displays. We collected nine icons that represented typical locations in cities, such as bus stations or museums. We created 150 fictional city maps using these stimuli. Finally, we selected a palette of colors different from those used on the park maps. The third and fourth columns of Table 1 list the RGB values for the colors used in the city map optimization tasks.

We assigned colors to the icons on the city maps to minimize search time, using the model that predicted the Experiment 1 data. We followed the same optimization procedure that we followed with the park maps.

The results of the optimization were similar to those obtained with the park map stimuli. When there were no restrictions on the number of colors that could be used, the optimization assigned each icon a unique color. When color assignments using more than four colors were penalized, the optimization produced displays that showed the same pattern of color assignments as obtained with the park maps. High-frequency search icons were assigned unique colors. Low-frequency icons were all assigned the same color. Figure 4C shows an example map. The number above each icon shows the search frequency for the icon. The number below each icon shows the predicted search time for that icon in milliseconds.

Figure 4D shows a color space distribution for the map. The three high-frequency icons were assigned colors distant from each other (mint green, aquamarine, and blue-gray). The optimization assigned the low-frequency icons fuchsia, a color in the lower-right quadrant of the color space distribution. Each low-frequency icon shares its color with six other icons. The optimal color for the low-frequency icons (search frequencies five, four, three, two, one, and zero) is a color that is distant from the other colors on the display. This pattern is identical to the one that was obtained for the optimized displays tested in Experiment 2.

The values in the fourth row of Table 4 show the average predicted search time for the low-frequency, medium-frequency, and high-frequency search icons. The model predicted that high-frequency icons would be found faster than either low-frequency icons or medium-frequency icons. We tested these predictions in Experiment 3.

## EXPERIMENT 3

The design of this experiment was similar to Experiment 2, except that the stimuli for this experiment were the 150 optimized city maps.

### Method

*Procedure*. For the third experiment, 15 observers participated. Participants searched for either a high-frequency, medium-frequency, or low-frequency target on each optimized display. The experiment followed the same procedure as Experiment 2. Observers from the Purdue University participant pool took part in the experiment and received course credit for their participation.

*Results*. Participants made incorrect responses on 3% of trials. The patterns of search times were similar to those obtained in Experiment 2. As search frequency increased, search time decreased, $F(2, 2031) = 104.12$, $p < .001$. The third row of Table 4 gives experimental search times for the low-, medium-, and high-frequency trials. The differences between search times attributed to search frequency indicate that the optimization assigned colors to make higher-frequency icons easier to find and produced measurable search time differences.

The model strongly predicted the search times for Experiment 3 ($R^2 = .31$, $p < .001$). This fit, however, was smaller than the one obtained for the predicted and experimental search times for Experiment 2 ($R^2 = .59$). The model's fit is worse because its parameters were set with the park map data from Experiment 1. As a result, the model does not predict search time for the Experiment 3 city maps as well as it did for the Experiment 2 park maps.

### Comparing Regression Models

It is interesting but may be difficult to compare the fit of a regression model whose parameters are set using the Experiment 3 data (the city map model) with the predictions of the model whose parameters were set using the Experiment 1 data (the park map model).

We performed linear regression using the search times from the 150 Experiment 3 trials as the dependent variable and color distinctiveness and eccentricity as factors. The effects of both factors were significant (color, $B_1 = -.27$, $t = -8.38$, $p < .001$; eccentricity, $B_2 = .55$, $t = 3.81$, $p < .001$). Color distinctiveness had a stronger effect in the city map model ($B_1 = -.27$) than in the park map model ($B_1 = -.18$), indicating that color has a stronger effect in the Experiment 3 setting. Eccentricity also had a stronger effect in Experiment 3 ($B_2 = .55$) than in Experiment 1 ($B_2 = .53$). The increase in the contributions of these factors is explained by the fact that the amount of variability in the target-distractor color distinctiveness scores and eccentricity scores is slightly higher for the city maps ($SD_{color} = 567.66$, $SD_{eccentricity} = 128.31$) than for the park maps ($SD_{color} = 549.50$, $SD_{eccentricity} = 113.65$).

The city map model fit the data with $R^2 = .36$. This fit is larger than the fit predicted by the park map model for those trials ($R^2 = .31$). This is expected, since the regression weights of the park map model were set with data from a visual search experiment that used different stimuli than the stimuli used in Experiment 3. However, the variance explained by the two models differs by only .05, indicating that the park map model was able to predict search performance on a visual search task with novel stimuli almost as well as the city map model, whose weights were set on the basis of the Experiment 3 data. This result supports the suggestion that the lower correlation between the park map model and the city map data largely reflects greater variability in the empirical data rather than poor parameter choices.

Although the city map model's fit for the Experiment 3 data ($R^2 = .36$) was worse than the park map model's fit of the Experiment 1 data ($R^2 = .47$), this finding should not be taken as evidence that the model is not generalizable. In fact, it is difficult to compare the fits of the two models directly, since the parameters of the models were selected to fit data from search experiments whose stimuli differed in important ways. The features of the Experiment 1 maps were designed to produce search times that varied as much as possible between maps. Thus, the colors for the icons on the Experiment 1 maps were assigned randomly, and each map contained up to 10 icon colors. In contrast, the Experiment 3 maps were optimized maps in which the icon colors were chosen to minimize predicted search time across icons without using more than four colors. The standard deviation of search times was lower in Experiment 3 (*SD* = 28 ms) than in Experiment 1 (*SD* = 173 ms). This finding suggests that the worse fit for the model in Experiment 3 may thus be attributable to a restricted range of search times in the optimized Experiment 3 stimuli rather than a limitation of the model.

## Optimizing for Other Characteristics

Once the search time model exists, it can be easily applied to assign colors to satisfy a variety of different design criteria.

## Minimizing Standard Deviation of Search Times

We have shown that it is possible to assign colors to promote fast search of frequent search targets. However, a designer may not know target search frequencies. Instead, a designer may want to choose the color assignment that makes each icon as easy to find as every other icon on the display. Accomplishing this task requires balancing the effects of eccentricity and target-distractor color distinctiveness. Icons near the display's center already have a search time advantage compared with icons near the display's edges. In addition, icons with unique colors are easier to find than icons that share colors with other icons. The color combination that equalizes search times across the possible targets is the predicted optimal assignment for that map.

*Optimization procedure*. We performed the four-color optimization on 150 map displays using a uniform search frequency distribution. For each display, the gradient-descent procedure was performed starting from 1,000 unique color assignments for each display. The final color assignment that produced the lowest predicted standard deviation of search times across icons was assigned to the display.

*Results*. Figure 6A shows an example of one of the optimized maps. The value under each icon is the predicted search time for that icon. The standard deviation of predicted search times for this map is 23 ms. For comparison, the standard deviation of fitted search times for the 450 Experiment 1 trials was 119 ms, indicating that the present optimization reduces variability of predicted search times. However, this comparison is not perfect, since we did not have participants search for multiple icons on each of the displays used in Experiment 1.

Icons at the edges of the screen have higher eccentricity values and thus longer predicted search times. The color assignment tries to counteract the eccentricity effect in two ways. First, the optimization assigns unique colors to icons at the edges of the display. Second, the optimization assigns colors to the icons at the edges of the display that are distant from the other colors on the display in color space. For example, Figure 6A contains a single yellow icon (the *restroom* icon in the upper right of the display). Uniquely colored icons usually have shorter search times than other display icons. However, yellow is close enough to orange in the color space distribution to reduce the yellow icon's color distinctiveness, yielding increased predicted search time for the yellow icon.

Icons at the center of the display have small eccentricity values and thus shorter predicted search times than icons at the display's edges. To bring the predicted search times of these icons closer to the predicted search times for the icons at the edges of the screen, the optimization assigns four of the central icons (dining, picnic, electricity, and cycling) the same color. Since more icons at the center of the map share the same color, predicted search time for these icons increases.

The results of the optimization to minimize standard deviation produced color assignments that reduced the predicted variability between search times for icons. The optimization approach is sophisticated enough to mediate between the effects of icon eccentricity and color distinctiveness factors in choosing optimal color assignments.

## Minimizing Search Time With a Uniform Search Frequency Distribution

The optimization that produced the maps tested in Experiment 2 involved a linear

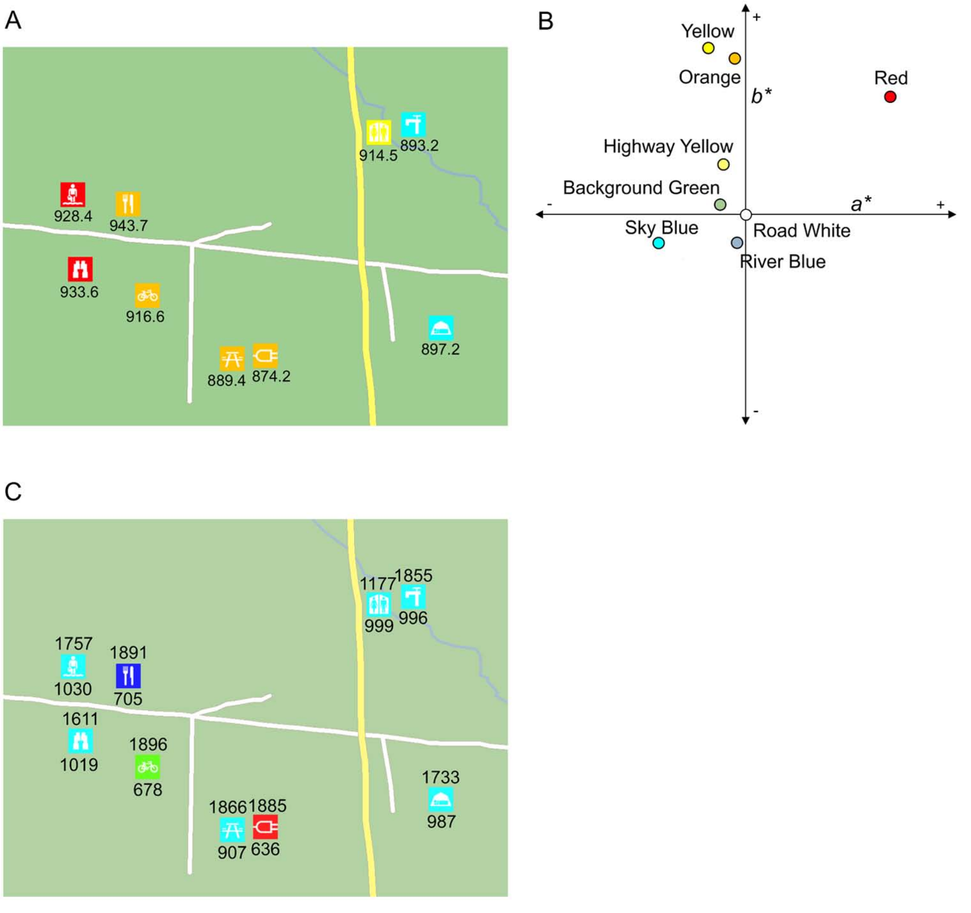


*Figure 6.* (A) Example optimized map from the optimization to minimize standard deviation. The value under each icon indicates predicted search time (in milliseconds) for that icon. The optimal color assignment balances the effects of target-distractor color distinctiveness and eccentricity. Icons at the edges of the screen were assigned distinctive colors. Icons near the center of the screen were assigned similar colors. Predicted standard deviation of search times is 23 ms. (B) Color histogram for the display in (A). (C) Example uniform search frequency optimized park map. All icons have equal search frequency. The number above each icon indicates the number of colored pixels that icon contains. The number below the icon indicates predicted search time (in milliseconds) for that icon. The three icons with the highest number of colored pixels were assigned unique colors. The rest of the icons were assigned the same color.

decreasing frequency distribution, whereby the most frequently searched icon was the target eight times, the second most frequently searched icon was the target seven times, and so on. The optimization's results reflected the influence of the frequency distribution: The optimization assigned unique colors to high-frequency icons at the expense of low-frequency icons. To examine the effects of the search frequency distribution on the color assignments the optimization chooses, we repeated the optimization using the four-color requirement but with a uniform frequency distribution. In this optimization task, we assumed that each icon would be the target an equal number of times, and the predicted search time for a color assignment was calculated as the average predicted search time across icons. The optimization chose four different colors that minimized the average predicted search time.

As before, we performed the optimization on 150 map displays. For each display, the optimization performed the gradient-descent algorithm starting from 1,000 unique color assignments. The final color assignment with the lowest predicted average search time was assigned to a display.

*Results*. When the goal was to minimize search time when all icons would be the target an equal number of times, the optimization assigned unique colors to the three icons with the highest number of colored pixels. The optimization assigned all other icons the same color. Figure 6C shows an optimized display. The number above each icon indicates the number of colored pixels the icon contains. The number below the icon indicates predicted search time (in milliseconds) for that icon.

As the figure shows, the three icons with the largest number of pixels were assigned unique colors. By assigning unique colors to the icons with the highest number of colored pixels, the algorithm reduces the number of colored pixels that would interfere with the other icons. This approach essentially removes the icons that would cause the most interference from the set of same-color distractors. On this display, the icons representing *dining, electricity*, and *biking* had the largest number of colored pixels, and the optimization assigned these icons unique colors.

These results illustrate that the shape of the search frequency distribution influences the strategy that the optimization uses to minimize average search time. When the frequency distribution decreases linearly, some icons will be searched for more often than others, so the optimization assigns those icons unique colors. However, when the frequency distribution is flat and every icon will be searched for equally often, the optimization assigns unique colors to the icons that have the most colored pixels.

## GENERAL DISCUSSION

The model fit the Experiment 1 data and was useful for choosing colors for icons on maps to minimize search time and minimize the standard deviation between predicted search times for icons. Experimental search times for the optimized maps obtained in Experiment 3 were correlated with the predictions of the search model, indicating that the optimization approach and the search model were effective.

We also showed that the model does not require additional data to predict search times for novel stimuli. We set the parameters of the model on the basis of data from Experiment 1, in which all displays had a green background. The novel search stimuli were displayed on a white and gray background. Color assignments for the two search situations reflected the properties of the backgrounds. Although the correlation between predicted and experimental search times was worse than the correlation between model and experimental search times for Experiment 1, the model's predictions did not drop simply to the correlations that would have been obtained if only eccentricity were considered. However, it is not known how much deviation between training stimuli and optimization stimuli the model can handle without breaking down.

When we began this project, we intended to use an already-existing model of visual search to model the data from Experiment 1. However, it soon became clear that no existing model would apply to the current task. First, as we noted in the Introduction, many models of visual search require representing a search display as a list of discrete objects in which all features are specified. These models cannot predict search time for natural images whereby the features of the display background may influence search time. In addition, few models of visual search that work with natural images apply to cases in which the observer searches for a known target. Our model thus represents a step in developing visual search models that work with natural images to predict search time for known targets.

The model we describe is relatively simple compared with other models of visual search. The model requires only one parameter in addition to the regression weights, the parameter that blurs the display color histogram in the calculation of target-distractor color distinctiveness (see the appendix for details). However, the model is flexible enough to make nuanced predictions that allow the optimization procedure to assign colors to minimize search time, which relies mainly on target-distractor color distinctiveness, and to minimize the standard deviation of search times, which requires balancing the effects of eccentricity and color distinctiveness.

Furthermore, the model makes no explicit assumption about observers' search strategies. That is, as discussed earlier, the model does not segment the search display into discrete objects that can be searched with the use of a particular strategy. Instead, we employ a measure of target-distractor distinctiveness that predicts the general difficulty of finding a specified target on a display. An additional motivation for avoiding explicit predictions about search strategy was that we did not instruct observers to use a particular strategy, and we did not measure eye movements. However, our measure of eccentricity captured a significant portion of the variability in search times related to the distance of a target from the center of the screen.

As we noted in the Introduction, several factors affect visual search. The current model contains representations of two of these factors. However, the model's fit to experimental data and its ability to choose optimal features for maps will almost certainly be improved by adding additional factors. For example, the model currently lacks representations of bottom-up factors known to influence search, such as the distinctiveness of an icon relative to its surroundings (Itti & Koch, 2000; Rosenholtz et al., 2007). In addition, the model's representation of a target's visual features is currently limited to color information. Although the current model can analyze the colors of targets with shapes other than squares, it does not represent shape information. Improving the model's representation of a target's visual similarity to distractors would improve its ability to fit search times for a wider variety of search stimuli (Alexander & Zelinsky, 2011).

If a designer knows that observers will employ a particular search strategy, the designer should certainly take advantage of that knowledge in choosing colors for icons on the display. For example, if a designer knows that observers tend to miss important information in one region of a display, the designer may want to use unique or high-contrast colors to draw attention to that region. Although it should be possible to extend our approach to include information about an observer's search strategy, our results show that the method we describe can produce color assignments that minimize search time even when this information is unknown.

Our approach in this article represents an effort to examine how quantitative models of human performance can inform design practice. In the current discussion, we have evaluated an empirical method for addressing the map design guideline that suggests choosing distinct hues for different categories when using a qualitative color scheme. Although other approaches (Campadelli et al., 1999; Carter & Carter, 1982) have identified sets of discriminable colors, these studies did not involve consideration of visual search processes and were not validated with data from observers. Thus, our approach represents an effort to seek empirical support for design principles that moves beyond prior work in this area.

It should be possible to integrate our approach with other design decisions involving color. As we noted in the Introduction, cartographic principles suggest taking advantage of existing color conventions when designing maps. Although none of the icon colors we considered in this article were used to convey meaning, the colors we chose for the map backgrounds used standard color code conventions (green = land, blue = water, etc.). Thus, the algorithm's selection of colors for the icons was constrained by other color choices that were guided by color conventions. In some cases, it may not be possible to consider alternative colors for some icons. For example, the colors of corporate logos may not be free to vary on a map. However, the method we describe should apply even when the colors of some items are fixed and the designer is trying to choose colors for icons or other display elements from a restricted color palette.

Imagine that a designer wants to assign colors to the icons in Figure 2. Color code conventions and standards suggest colors for at least some of these icons. The wheelchair symbol, also known as the "international symbol of access," should be blue (Ben-Moshe & Powell, 2007). Also, four of the icons have to do with water or nature (the *faucet* icon, the *fishing* icon, the *shower* icon, and the *boating* icon), so the designer might choose different shades of blue or green for them as well. Depending on the context, the designer might make the *electricity* icon red to indicate danger. However, this color assignment still leaves three icons (bathroom, dining, and parking) with no obvious color code conventions. If the designer has a list of possible candidate colors for these icons, our approach can suggest color choices that will make the remaining icons stand out from the other icons on the display (by minimizing average search time for those icons relative to the other display icons) or make them as easy to find as the other icons (by minimizing standard deviation of search times across icons), depending on the designer's preference.

## CONCLUSIONS

We used a model of visual search to choose colors for icons on maps to minimize either average search time or standard deviation of search times. We examined the model's ability to assign colors in a variety of optimization tasks. Experimental and predicted search times for the optimized maps were significantly correlated, indicating that the optimization approach was effective.

The model's ability to choose colors for other map displays will likely depend on the similarity of the new displays to the Experiment 1 maps used to set the model parameters. We have shown that the model is somewhat flexible, since it chose optimal color sets for the Experiment 3 maps, which used different icons, colors, and backgrounds than did the Experiment 1 maps. However, a designer who wants to choose optimal colors for stimuli that are strongly different from the Experiment 1 stimuli may need to choose new model parameters appropriate to the new search situation. Doing so would require collecting search time data from a visual search experiment with the new stimuli and setting the parameters of the model with the new data.

A visual search model able to make accurate quantitative predictions about what displays will lead to faster searches for specific icons would be a useful tool in an interface designer's toolbox, since it could suggest colors for icons when no color code conventions exist. The visual search model and optimization approach reported here represent a useful step toward developing such a tool. In future work, we will examine the model's ability to choose colors for single icons when other display features are held constant and will compare the search time benefits of following this approach with the benefits of following color code conventions.

## APPENDIX

### Histogram Backprojection

Histogram backprojection (Swain & Ballard, 1991) is used to calculate target-distractor distinctiveness scores for color in our model. First, each feature of a search target is encoded in a histogram whose bins represent the range of values of that feature. To construct the search target histogram, *T*, the algorithm examines the search target's pixels one by one and calculates the total number of pixels that fall into each feature bin. A display histogram, *D*, is produced by the same method. In addition, a reference matrix, *R*, is created that has the same dimensions as the target image. Each cell in the reference matrix contains the indices of the bin in the target histogram *T* that the corresponding pixel in the target incremented.

For example, in the simplified trial shown in Figure A1, the target image contains one red item. The red item contributes to the total in the red bin of the target histogram. The value in the reference matrix is 1, the index of the red bin.

Next, the target histogram is divided bin by bin by the display histogram to produce a co-occurrence histogram, *C*:

$$C_i = \frac{T_i}{D_i}, \tag{A1}$$

where $T_i$ is the number of pixels in bin *i* of the target histogram, $D_i$ is the number of pixels in bin *i* of the display histogram, and $C_i$ is the number of pixels in bin *i* of the co-occurrence histogram. Each bin in *C* has values between 0 and 1 and represents the amount of each of the target's features that the display contains. If the display contains many pixels that have the same feature as the target, the value of *C* for that

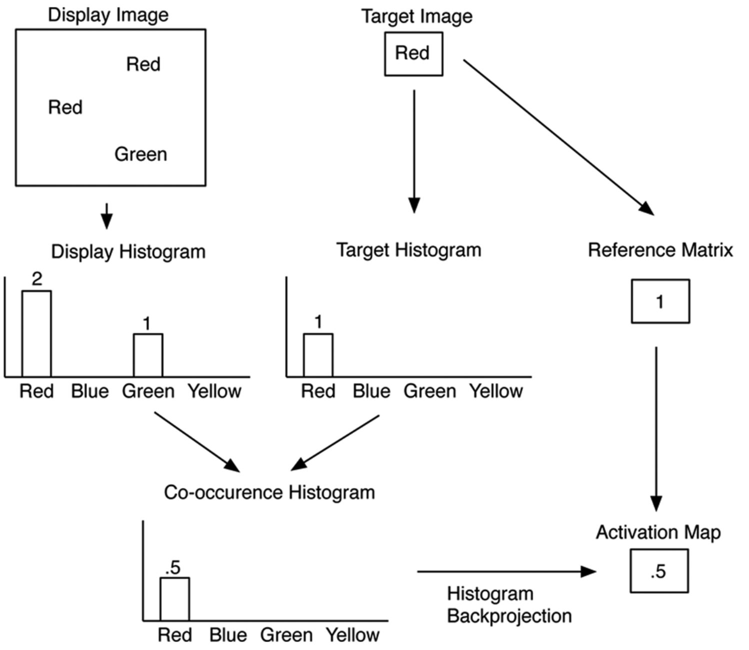


*Figure A1.* The histogram backprojection procedure when the target is a red item and the display contains red and green items. Target-distractor distinctiveness is 0.5.

feature will be close to 0. When the display contains no pixels that have the same feature as the target (other than the pixels that belong to the target), the value of $C$ for that feature will equal 1.

To calculate a final target-distractor distinctiveness score for a particular target on a display, the co-occurrence values for each feature must be weighted in terms of how often those features occur in the target. The histogram backprojection procedure accomplishes this by translating or "backprojecting" the values in the co-occurrence histogram into an activation map, $A$, that has the same dimensions as the target. The activation map is created by using the values in $R$ as lookup coordinates in the co-occurrence histogram $C$:

$$A_{x,y} = \mathrm{C}_{R_{x,y}} \tag{A2}$$

where $x$, $y$ is a location on an activation map $A$ and $\mathrm{C}_{R_{x,y}}$ is a lookup value in the co-occurrence matrix indexed by the corresponding value in the reference matrix $R$. Each value in $A$ indicates the distinctiveness of the feature located at each pixel in the target, relative to the number of pixels in the display that share that feature. When a target feature is unique, the value at each location in the activation map corresponding to that feature is close to 1. When distractor icons share a feature with the target, activation is reduced at each location in the activation map corresponding to that feature.

Summing the values of the activation map

$$TD = \sum_x \sum_y A_{x,y} \tag{A3}$$

gives us a measure of target-distractor distinctiveness, $TD$, for a given feature on a trial.

Figure A1 shows an example of the backprojection procedure when the target is a red item and the display contains a distracting red item and a green item. First, target and display histograms are created. In this simplified example, the color histogram contains bins for four colors. The count in each bin represents the number of icons of each color. The right column of the figure shows the reference matrix that contains the index of each histogram bin that the color of the target incremented in the target histogram. After the histograms have been created, the target histogram is divided bin by bin by the display histogram. Since the target counts as one red item and there are two red items on the display, the value in the co-occurrence histogram for red is 1/2 or 0.5. The lower portion of the figure shows the histogram backprojection step. The values in the bins of the co-occurrence histogram are backprojected into the activation map. Finally, the values in the activation map are summed across display locations to obtain the target distractor distinctiveness score.

## Target Distractor Color Distinctiveness

To calculate the target-distractor color distinctiveness for a trial, we create color image histograms for the target and display images. First, we transform RGB values into CIELAB color values (Ruzon, 1998; Fairchild, 1998). The CIELAB color space represents differences in color appearance as distances in a three-dimensional space. The $L*$ dimension of the color space contains luminance information, and the $a*$ and $b*$ dimensions contain color information. For simplicity and computational efficiency, we ignore the $L*$ dimension. After converting the target and display color values, we create a 240 × 240 bin histogram for each image, with axes representing the $a*$ and $b*$ dimensions of the CIELAB space.

In the CIELAB color space, values in the $a*$ and $b*$ dimensions are in the range −120 to 120. To create a color histogram for an image, we first rescale the values in both the $a*$ and $b*$ dimensions of the image using the formulas

$$a_r = \text{round}\{[(a^* + n_{\text{bins}} / 2) / n_{\text{bins}}] \times (n_{\text{bins}} - 1)\} + 1 \text{ and} \quad \text{(A4)}$$

$$b_r = \text{round}\{[(b^* + n_{bins} / 2) / n_{bins}] \times (n_{bins} - 1)\} + 1, \quad \text{(A5)}$$

where $n_{\text{bins}} = 240$ and $a_r$ and $b_r$ are the rescaled CIELAB $a*$ and $b*$ values. The $a_r$ and $b_r$ values are integers in the range 1 to 240 that correspond to the bins in a 240 × 240 histogram. The color histogram is created by moving pixel by pixel through the rescaled image and incrementing the value in the histogram bin that corresponds to each $(a_r, b_r)$ value. Each bin in the histogram gives the number of pixels in the image of a certain color.

Including antialiased pixels in the color histograms adds unwanted noise to the calculation of color distinctiveness, so we limit the color information that can contribute to the color histogram. Only the nonantialiased color pixels and white pixels in the target and display images contribute to the color histograms for those images.

The histogram backprojection approach described by Swain and Ballard (1991) involves consideration of differences between neighboring colors in color space as all or none. In terms of predicting search time, histogram backprojection predicts that the distinctiveness of a target color is independent of its similarity to other colors on the display. In other words, the model predicts that a red icon among green icons is as distinctive as a red icon among nearly red icons, as long as the nearly red icons are in a different bin than the red icon.

Since this approach does not model human behavior, we include a measure of color distance in the model so that the color distinctiveness score for a target is reduced when a display contains icons with colors similar to the target. After creating the target and display histograms, the model blurs the display histogram using a 71 × 71 0th-order Gaussian filter:

$$G(x, y) = e^{-(x^2 + y^2) / 2\sigma^2}, \quad \text{(A6)}$$

where $x = [–35 : 35]$, $y = [–35 : 35]$, and $\sigma = 8$. We considered $\sigma$ values between 0 and 20. Setting $\sigma = 8$ produced the best observed fit between predicted search times and experimental search times (described later). Figure 1C shows the size of the Gaussian filter relative to the map displays. When the display histogram is blurred during the calculation of target-distractor color distinctiveness with use of the Gaussian filter, the number of pixels in the bins for colors near other colors in color space is increased, reflecting the similarity of the colors.

An icon's target-distractor color distinctiveness score is influenced by the number of colored pixels in the icon. As Figure 6C shows, the number of colored pixels in an icon differs across the icons used in Experiment 1. The value above each icon in Figure 6C indicates the number of colored pixels in an icon.

Imagine a case in which the target's color is unique on the display. $C_i$ for the bin corresponding to the target's color will equal 1, since only the pixels in the target contribute to the bin in the display histogram corresponding to the target's color. When the co-occurrence histogram is backprojected into the activation map *A*, every colored pixel of the target will get a value of 1. Considering only the colored pixels of the target, one finds that the *TD* value for color will be the sum across the values on the activation map. As a result, targets with more colored pixels will receive higher target-distractor color distinctiveness scores and shorter predicted search times than will targets with fewer colored pixels.

Adding the $L^*$ dimension to the color image histogram we used to calculate target-distractor distinctiveness would probably have improved the fit of the model. However, it would have increased the size of each target and display color histogram by a factor of 240. The target and display histograms used in the current project contained 57,600 bins (240 for the $a^*$ dimension × 240 for the $b^*$ dimension). Adding an $L^*$ dimension would have resulted in histograms containing 13.8 million bins ($240^3$).

Increasing the sizes of the target and display histograms would have made the optimizations more time-consuming than they already were. The step in histogram backprojection that requires the most time is the histogram blurring. If the histograms contained three dimensions, this blurring would have taken dramatically longer, thus increasing the time required to evaluate each of the possible assignments of colors during the optimizations. For these reasons, we did not include the $L^*$ dimension.

## KEY POINTS

- Search time data from a map search experiment can be modeled with measures of target-distractor color distinctiveness and target eccentricity.
- Histogram backprojection provides an effective method for operationalizing target-distractor color distinctiveness.
- Gradient-descent algorithms can be used to choose color assignments to minimize predicted search time and predicted standard deviation of search times across icons.
- The optimal color for a display element depends on the relative importance of the element.
- Quantitative models of search behavior can inform design choices.

## REFERENCES

Alexander, R. G., & Zelinsky, G. J. (2011). Visual similarity effects in categorical search. *Journal of Vision, 11*, 1–15.

Ben-Moshe, L., & Powell, J. J. W. (2007). Sign of our times? Revis(it)ing the international symbol of access. *Disability & Society, 22*, 489–505.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Bravo, M. J., & Farid, H. (2008). A scale invariant measure of clutter. *Journal of Vision, 8*, 1–9.

Brewer, C. A. (2005). *Designing better maps: A guide for GIS users*. Redlands, CA: ESRI.

Campadelli, P., Posenato, R., & Schettini, R. (1999). An algorithm for the selection of high-contrast color sets. *Color Research and Application, 24*, 132–138.

Carter, R. C., & Carter, E. C. (1982). High-contrast sets of colors. *Applied Optics, 21*, 2936–2939.

Dent, B. D. (1990). *Cartography.* Dubuque, IA: Wm. C. Brown.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433–458.

Duncan, J., & Humphreys, G. W. (1990). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 578–588.

Fairchild, M. D. (1998). *Color appearance models*. Reading, MA: Addison-Wesley.

Francis, G., Bias, K., & Shive, J. (2010). The psychological four-color mapping problem. *Journal of Experimental Psychology: Applied, 16*, 109–123.

Grossberg, S., Mingolla, E., & Ross, W. (1994). A neural theory of attentive visual search: Interactions of boundary, surface, spatial, and object representations. *Psychological Review, 101*, 470–489.

Harrower, M., & Brewer, C. A. (2003). ColorBrewer.org: An online tool for selecting color schemes for maps. *Cartographic Journal, 40*, 27–37.

Harvey, F. J. (2008). *A primer of GIS: Fundamental geographic and cartographic concepts*. New York, NY: Guilford Press.

Hwang, A. D., Higgins, E. C., & Pomplun, M. (2009). A model of top-down attentional control during visual search in complex scenes. *Journal of Vision, 9*, 1–18.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489–1506.

Kraak, M. J., & Ormeling, F. J. (1996). *Cartography: Visualization of spatial data*. Harlow, UK: Longman.

Liu, B., Francis, G., & Salvendy, G. (2002). Applying models of visual search to menu design. *International Journal of Human-Computer Studies, 56*, 307–330.

Najemnik, J., & Geisler, W. (2005). Optimal eye movement strategies in visual search. *Nature, 434*, 387–391.

Navalpakkam, V., & Itti, L. (2006). Top-down attention selection is fine-grained. *Journal of Vision, 6*, 1180–11931.

Navalpakkam, V., & Itti, L. (2007). Search goal tunes visual features optimally. *Neuron, 53*, 605–617.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Phillips, R. J. (1979). Making maps easy to read: a summary of research. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing of visible language I* (pp. 165–174). New York, NY: Plenum.

Rosenholtz, R. (2001). Search asymmetries? What search asymmetries? *Perception & Psychophysics, 63*, 476–489.

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision, 7*, 1–22.

Ruzon, M. (1998). RGB and CIE Lab converters [Computer software]. Palo Alto, CA: Stanford University. Retrieved from http://robotics.stanford.edu/ruzon/software

Scialfa, C. T., & Joffe, K. M. (1998). Response times and eye movements in feature and conjunction search as a function of target eccentricity. *Perception & Psychophysics, 60*, 1067–1082.

Shive, J., & Francis, G. (2008). Applying models of visual search to map display design. *International Journal of Human-Computer Studies, 66*, 67–77.

Swain, M. J., & Ballard, D. H. (1991). Color indexing. *International Journal of Computer Vision, 7*, 1573–1405.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Van Laar, D., & Deshe, O. (2002). Evaluation of a visual layering methodology for colour coding control room displays. *Applied Ergonomics, 33*, 371–377.

Williams, D. E., Reingold, E. M., Moscovitch, M., & Behrmann, M. C. (1997). Patterns of eye movements during parallel and serial visual search tasks. *Canadian Journal of Experimental Psychology, 51*, 151–164.

Williams, L. G. (1967). The effects of target specification on objects fixated during visual search. *Acta Psychologica, 27*, 353–360.

Wolfe, J. M. (2007). Guided Search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems* (pp. 99–119). New York, NY: Oxford.

Yantis, S. (1993). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24*, 295–340.

Yeh, M., & Wickens, C. D. (2001). Attentional filtering in the design of electronic map displays: A comparison of color coding, intensity coding, and decluttering techniques. *Human Factors, 43*, 543–562.

Zelinsky, G. (1997). Using eye saccades to assess the selectivity of search movements. *Vision Research, 36*, 2177–2187.

Zelinsky, G. (2008). A theory of eye movements during target acquisition. *Psychological Review, 115*, 787–835.

Joshua Shive is an assistant professor of psychology at Tennessee State University in Nashville. He earned his PhD in cognitive psychology from Purdue University in 2008.

Gregory Francis is a professor of psychological sciences at Purdue University in West Lafayette, Indiana. He earned his PhD in cognitive and neural systems from Boston University in 1994.

*Date received: December 6, 2010*
*Date accepted: July 16, 2012*