Framework to analyze the use of colors in informational maps related to Covid-19 data

Sara Copetti Klohn, Anelise Zimmermann

This paper presents a study about using colors in maps that present information related to the Covid-19 pandemic. A sample was collected from two different governmental websites of the state of Rio Grande do Sul, Brazil. An analysis framework was developed based on the sample and drew from a literature review about the use of colors and maps in data visualization. The analysis showed that one of the examples collected presents an inconsistent use of colors, which could prompt difficulties in visualizing the data; the other example uses color more consistently, potentially being more effective in communicating information. In addition, the study showed potential for continuity and expansion due to the relevance of information design applied to health information.

1 Introduction

Complex medical and health information can be simplified with uncomplicated language, presented with visual features to reach audiences with different literacy levels. For example, a sequence of illustrations can demonstrate procedures (Spinillo, 2000; Spinillo, Oliveira, Mazza, Castro Lima, & Assis, 2020); infographics can explain the vaccine approval process, and maps can show a significant amount of data about confirmed cases of a given disease throughout a territory. These examples have flooded the news and social media in the past year due to the ongoing Covid-19 pandemic, to date spanning 2020 and 2021. The role of information design in conveying medical and scientific information to the general public has never, in recent history, been so prominent around the whole world. However, this sudden increase in data visualization does not necessarily mean that people are taking in more information or becoming better informed. For instance, excess data with no apparent visual parameters can lead to misunderstanding and misconception. The different aspects that can influence the comprehensibility of visual information such as colors, shapes, symbols, pictures have been extensively investigated (Kostelnick, 2016; Lonsdale et al., 2019; Lyra, Reis, Cruz, & Isotani, 2019; Mayer, 2009; Schnotz, Picard, & Hron, 1993). Hence, this study focuses on the use of colors.
In this study, we specifically look at the use of colors in maps that show causalities and spread of the coronavirus on governmental websites of Rio Grande do Sul (RS), Brazil’s southernmost state. The analysis focused on the use of colors in maps that present the spread of the virus, bearing in mind two key points: 1. the cognitive load that visuals create in the reader (Mayer, 2009) and 2. that the lack of obvious functionality of specific design features can create unnecessary interpretative demands on the reader (Kostelnick, 2016). The samples were collected on two different occasions, in February and March of 2021. A framework developed based on theoretical research about colors in data visualization was used to run the analysis. To conclude, analysis outcomes are discussed as well as suggestions for the study’s continuity and expansion.

2 Maps in visual information

As much as helping scientists, governments, and health organizations to understand and control the virus, the visualization of Covid-19’s spread and casualties has been used to inform the public about the situation and risk in their area. Visual features that make the information accessible to audiences give them confidence to explore data, empowering decision-making, even when facing anxiety-inducing risks (Kostelnick, 2016). In addition, clear, concise, and user-friendly maps translate complex content to audiences in readable and understandable ways (Cicalò & Valentino, 2019).

Alongside other medical data, geographic coordinates allow researchers to identify a disease’s origins, spread, and tendencies (Caprarelli & Fletcher, 2014). Mapping diseases has been used to analyze outbreaks starting form at least the 17th century (Cicalò & Valentino, 2019); John Snow’s dot map of cholera occurrence is famous for demonstrating the connection of polluted water and the high number of cholera cases in the 1854 epidemic (Pettersson, 2015). Besides serving as a research tool, mapping diseases is a valuable way of spreading information and engaging the public to take the appropriate preventive measures.

Maps are infographic devices, according to Rajamanickam (2005), that can be used as (i) locators, (ii) data collection, and (iii) schematic representation of a surface or process. In the first usage, they show the location of one point in comparison to the position of another point. Second, they relate quantitative data distributed through a territory. Finally, they simplify and delineate the representation of space, process, and sequence in an abstract, conceptual, and bidimensional way.

There are some aspects concerning the type of information and how it is presented that need to be considered when showing sensitive data to the general public. Real-time data about a disease’s spread presented in maps, for example, might elevate the audience’s apprehension and fear. “The dynamism of temporal proximity both
energizes and enervates readers, depending on the corresponding emotions it engenders, ranging from anxiety or expectation to joy or euphoria, disappointment or dejection.” (Kostelnick, 2016, p. 128). On the other hand, “because of the proximity of the danger that threatens them or their loved ones, audiences become engaged emotionally in the visualization.” (Ibid., p. 128). Weber (2017), states that infographics, mainly those based on maps, can be misleading, as they are models and do not mirror concrete objects; they always pass through the designer’s subjectivity and interpretation in terms of its visual features. Hence, visual information needs to be adequately designed, i.e., using visual features in a meaningful, cohesive, and clear manner, considering all visual aspects involved.

Lyra et al. (2019), based on cognitive load theory (intrinsic, extraneous, and germane loads), established levels of complexity of infographics based on cognitive demands of the features depicted. For example, a map legend’s necessity is dependent upon the type of information on a map, combined with the reader’s familiarity with visual language. Typically, map legends indicate places, represent a quantitative hierarchy, and establish associations between information. According to Lyra and colleagues, if the reader has to memorize the associations between information and its visual presentation without the help of a legend, an infographic’s content is considered to have high complexity. On the other hand, if the visual information and codes presented are well known by the audience, the necessity to memorize associations to read and understand the information is reduced; thus, the content is considered to have low complexity. In a nutshell, the higher the complexity level, the higher the cognitive load and therefore more difficult to understand.

On maps’ legibility, Pettersson (2021) points out some visual variables that should be observed, which include: contrast in dimension, contrast in form, the complexity of patterns (for instance, granularity or texture), number of typefaces, bold and distinct symbols in a consistent size, position and place, directions, and, finally color, which also include density or greyness. Considering all these aspects, the ability to read and interpret a map will depend in part on the reader’s visual literacy experience and in part on the contrast, consistency, and the number of variable features. While contrast is fundamental to discern information, consistency of features guides the reader in reading the map. Therefore, if the variability of features is relevant to create a contrast among the visual elements, a significant number of variances can suggest a lack of consistency.

The roles of colors on maps will be described below.

3 Colors in visual information

Color is a feature of visual information that can contribute or disrupt communication. For instance, “the use of color on maps introduces a
large number of variables, which may enhance contrast, and extend the number of perceptual differences that can be employed in discrimination.” (Pettersson, 2021, p. 30). Conversely, color variability also augments the number of features to be “read” in the artifact, possibly increasing the chance of misunderstanding or lack of a clear association. Furthermore, color creates meaning that is culturally constructed (Kostelnick, 2016), such as red or yellow for danger or attention, coded as traffic lights and signs, as well as physical aspects, like blue for water and green for vegetation.

Deploying unfamiliar color coding can make the information more complex, and, therefore, more difficult to understand. This is because the reader cannot mentally process too much information at once, visual or written. Reducing information to what is essential also reduces the cognitive load placed upon the reader, which aids in comprehending information. In this scenario, presenting information in chunks can reduce the cognitive load (Few, 2004; Lonsdale et al., 2019; Lyra et al., 2019). Having too many colors for different sets of data forces the reader to go back and forth between the chart and legend, overloading their working memory, which cannot take in more than three or four chunks of information at a time (Lonsdale et al. 2019). Likewise, using a gradient of colors to show numeric data variation in a chart is confusing, as the reader cannot interpret numbers in a continuum of color ranging from, for example, red to blue (Few, 2005).

In an informational artifact, different categories should be able to be quickly identified by the viewer, and Pettersson (2017) argues that Gestalt principles of similarity and contrast are fundamental for this. Consistent use of color and typography to categorize information shows similarity. Contrast is important because the reader only perceives differences through comparison: small only exists in comparison with large. Likewise, something is old when opposite of something newer; if the same “old” thing is opposite of something older, then it would be seen as new (Cardoso, 2017).

Menezes and Pereira (2017) classified three categories for color purpose in infographics, based on constructions of meaning and people’s perception. The first category is the perceptive function, which refers to what people see without interpretation of meaning. Here, color is used to attract, harmonize, organize, provide visibility and readability. The second, indicative function, relates to the use of color to guide the reader through the artifact, similar grouping of information and help the reader in constructing the meaning; color is used to label, measure, rank, and maintain consistency. The last and third category is representative function and describes color that helps identify ideas and objects through similarity and symbolizing real objects or cultural conventions.

As such, visual conventions are commonly used to communicate messages rapidly. These conventions are cultural and social, and have local meaning and should be shared by a community of
users (Kostelnick, 2017). Color conventions are typically used to communicate culturally conveyed meaning (such as green for permission and red for prohibition) as well as to emphasize (e.g., bright contrasting colors) or mute information (e.g., greyscales or muted colors) (Weber, 2017; Pettersson, 2017). Essentially, to emphasize the core information, it is necessary to reduce the number of details to what is essential, enhancing attention and perception (Pettersson, 2015, 2017). Highlighting relevant information (in pictorial warnings) was found to increase comprehension, while highlighting less relevant information can reduce pictorial comprehension (McDougald & Wogalter, 2014). Therefore, colors can indicate information hierarchy. Ideally, soft colors should be used as a base for visualizing data and bright colors to highlight important information, as persistent use of bright colors accosts visual perception (Few, 2005, 2008).

Lonsdale et al. (2019, pp. 45–46) gathered design principles for online information, infographics, and motion graphics. In their paper, they present a table based on work from leading authors that shows guidelines for using visual features such as typography, colors, and pictograms. From their table, we selected the recommendations for colors in online information and infographics, as shown in Chart 1 below.

Next, bearing in mind the theory presented in this section, we discuss the methods and the framework developed for analyzing the sample collected.

**Chart 1**  Principles for using colors in online information and infographics, extracted from Lonsdale et al.’s (2019, pp 45–46) table: ‘Design principles framework used to inform design solutions to the problem identified in this study’.

| **Online Information** | Color should be used sparingly, with good contrast for text and images, and as an information tool (not as decoration).  
Color can be used to influence user satisfaction and trust. Users seem to prefer the colors blue and orange on a website, with the presence of orange increasing information recall (probably due to increased attention).  
The cultural influence of a color’s perceived meaning should be considered.  
Color coding should be used in a consistent and logical manner and should be understood quickly.  
Larger areas of white space should be used on a webpage to increase clarity and prevent the appearance of cluttered information. |
| **Infographics** | Colors must reflect the subject matter and fulfill specific needs and purposes.  
A color palette should be, on average, between 3-5 colors.  
Color can be used to help group chunks of information, to emphasize certain words, show hierarchy and relationships between elements, to help navigate the information.  
Color coding can be used to show levels of severity. |
4 Methods for the analysis

The data analyzed was collected from RS’s official governmental websites, which use the state’s map to inform the general public about the spread of Covid-19 in the area to date. They also include information concerning contagion rates, hospital occupation, and other related information. The sample was collected in February 2021 by taking screenshots of what people see when looking at a given map at a given time.

Two maps that are representative of the visual information displayed on the websites were selected and can be seen in Figures 1 and 2 below. Example 1 (Figure 1) depicts confirmed cases of Covid-19 on the side of a coronavirus panel, which displays other information such as the number of hospitalizations, and Intensive Care Unit (ICU) occupancy. Example 2 (Figure 2) shows the flag system developed by RS’s government that indicates levels of social distancing necessary to guide actions the public needs to take.

For the analysis, a framework was developed based on the theoretical review already presented, and the sample collected. This type of observation creates an organic and fluid process, responsive to the sample, to determine categories of analysis (Klohn, 2018). Frameworks are a common tool to analyze information design artifacts (Dyson, 2017) as they are sensitive to issues emerging from the material analyzed and consider pre-established theories (Stahl-Timmins, 2017). In this study, the analysis was made independently by the two authors, using the framework. The results were then compared and discussed to check for any dissonance over the outcomes.

Figure 1 Map 1 – confirmed Covid-19 cases in Rio Grande do Sul, on the coronavirus panel section of the website. Source: https://ti.saude.rs.gov.br/covid19/
The framework designed is shown in Chart 2 and outlines nine categories (first column); each one is classified according to Menezes and Pereira’s (2017) definition of color functions (second column).

### Chart 2 Framework of analysis.

<table>
<thead>
<tr>
<th>Category of analysis</th>
<th>Color purpose</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) Use of white space areas to declutter information</td>
<td>Perceptive</td>
<td>Lonsdale et al. (2019)</td>
</tr>
<tr>
<td>(B) Reduced number of colors</td>
<td>Perceptive</td>
<td>Lonsdale et al. (2019), Few (2005)</td>
</tr>
<tr>
<td>(C) Highlights core information</td>
<td>Indicative</td>
<td>Weber, 2017; McDougald &amp; Wogalter (2014); Um et al. (2012)</td>
</tr>
<tr>
<td>(D) Bright colors for highlights</td>
<td>Indicative</td>
<td>Few (2005)</td>
</tr>
<tr>
<td>(E) Consistent use of color according to Gestalt principle of similarity</td>
<td>Indicative</td>
<td>Petterson, 2017</td>
</tr>
<tr>
<td>(F) Consistent use of color according to Gestalt principle of contrast</td>
<td>Indicative</td>
<td>Pettersson, 2017; Cardoso, 2017; Lonsdale (2019)</td>
</tr>
<tr>
<td>(G) Legend for unknown visual codes</td>
<td>Indicative</td>
<td>Lyra et al. (2019)</td>
</tr>
<tr>
<td>(H) Use of clear color conventions</td>
<td>Representative</td>
<td>Kostelnick (2017); Pettersson (2017); Weber (2017)</td>
</tr>
<tr>
<td>(I) Color coding to show levels of severity</td>
<td>Representative</td>
<td>Lonsdale et al. (2019)</td>
</tr>
</tbody>
</table>

**Figure 2** Map 2 – regions flag system that indicates levels of social distance due to Covid-19 cases in Rio Grande do Sul. Source: https://distanciamentocontrolado.rs.gov.br
There are two perceptive categories (A, B), which are related to the roles of harmonizing, organizing, and facilitating visibility and legibility; five of the categories are indicative (C, D, E, F, G) with the purpose of labeling, measuring, creating hierarchy, and maintaining consistency; and two categories are representative (H, I) as they identify and symbolize information. Finally, the third column of the framework indicates the theoretical basis for each category.

5 Analysis of Covid-19 data maps in Rio Grande do Sul

First, the framework was used to determine if the two selected maps matched each category recommendation with a simple answer of yes, no, or partially by each researcher. The categories were then checked for their clarity and any overlaps among them, resulting in the chart below.

The sample colors were analyzed considering all information within the section of the webpage where they were placed. This decision was made because in Example 1, although there is a clear vertical section division, there is no clear horizontal section division, which means both sides can be read together (Zone 3 and 4 marked on Figure 3 below). Charts 3 and 4 below show the analysis of Examples 1 and 2.

The analysis identified areas where the use of color is pertinent, and others where there are potential risks of misunderstanding. In both maps, successfully, shades ranging from yellow to red demonstrate levels of severity (category I), which is a well-known code that resembles traffic lights systems. Less successfully, in Example 1, there is not a clear legend identifying color codes (G).

Next, the positive and negative aspects of the sample are discussed in more detail and each example was split into zones to facilitate the discussion (Figures 3 and 4 in the following section).

<table>
<thead>
<tr>
<th>Category of analysis</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) Use of white space areas to declutter information</td>
<td>Partially</td>
</tr>
<tr>
<td>(B) Reduced number of colors</td>
<td>No</td>
</tr>
<tr>
<td>(C) Highlights core information</td>
<td>Partially</td>
</tr>
<tr>
<td>(D) Bright colors for highlights</td>
<td>No</td>
</tr>
<tr>
<td>(E) Consistent use of color according to Gestalt principle of similarity</td>
<td>Partially</td>
</tr>
<tr>
<td>(F) Consistent use of color according to Gestalt principle of contrast</td>
<td>Partially</td>
</tr>
<tr>
<td>(G) Legend for unknown visual codes</td>
<td>Partially</td>
</tr>
<tr>
<td>(H) Use of clear color conventions</td>
<td>Partially</td>
</tr>
<tr>
<td>(I) Color coding to show levels of severity</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Chart 3 Framework analysis of Example 1 (Figure 1).
5.1 Example 1 analysis

The first map shows the number of Covid-19 confirmed cases through a color-coded system. As mentioned, there is no clear division between Zone 3 (Z3) and 4 (Z4) on Example 1 (Figure 3), and both sides of this webpage can be seen as one piece of information. However, the chart on the left (Z3) does not directly relate to the map on the right (Z4). Although the colors do somewhat contribute to the understating of the information, mainly they disrupt it.

The similar colors (red and yellow) on the map (Z4) and on the charts (Z3) do not correspond informationally. The map ranges from

---

**Chart 4** Framework analysis of Example 2 (Figure 2).

<table>
<thead>
<tr>
<th>Category of analysis</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) Use of white space areas to declutter information</td>
<td>Yes</td>
</tr>
<tr>
<td>(B) Reduced number of colors</td>
<td>Yes</td>
</tr>
<tr>
<td>(C) Highlights core information</td>
<td>Yes</td>
</tr>
<tr>
<td>(D) Bright colors for highlights</td>
<td>Yes</td>
</tr>
<tr>
<td>(E) Consistent use of color according to Gestalt principle of similarity</td>
<td>Yes</td>
</tr>
<tr>
<td>(F) Consistent use of color according to Gestalt principle of contrast</td>
<td>No</td>
</tr>
<tr>
<td>(G) Legend for unknown visual codes</td>
<td>Yes</td>
</tr>
<tr>
<td>(H) Use of clear color conventions</td>
<td>Partially</td>
</tr>
<tr>
<td>(I) Color coding to show levels of severity</td>
<td>Yes</td>
</tr>
</tbody>
</table>

---

**Figure 3** Division of Example 1 in four areas for discussion.
light yellow to red, representing the number of confirmed cases of Covid-19, while the charts show yellow for hospitalizations and red for intensive care occupation and, separately, to highlight the number of deaths and apparent virus lethality. Adding to the lack of clear color convention, this aspect also suggests the need for more white (or neutral color) areas declutter information (A).

The large variety of colors contradicts the recommendation to use fewer colors (B) to reduce the reader’s cognitive demands. Besides the shades of red and yellow representing levels of severity, the colors of the landscape features (physical), and the background of Z3, most colors are considerably bright and vivid, which makes it challenging to identify the page’s core information (C and D). Therefore, the contrast (F) between colors does not have a distinguishable meaning and does not assist in comprehending the information.

Finally, the legend (G) next to the map identifies the code for the shades of yellow to red, indicating the number of people infected by the virus in the state. Although the legend is clear, and the similarity (E) of the shades of color makes sense, using well-known codes for severity levels, the subtle variation between the eight shades might increase difficulty in reading the information. In addition, the constant need to go back and forth between map and legend creates extra cognitive demands on the reader (Lyra et al., 2019). Furthermore, color shades might not make it easy for the reader to compare the legend to regions, especially considering the several subdivisions of the state’s regions.

5.2 Example 2 analysis

The map on Example 2 relates to a flag system that identifies infection risk levels throughout the state. This example has more efficient use of colors than Example 1, although part of the information necessary to understand the meaning of the colors is hidden in a drop-down menu on the webpage. There sections are also divided better than the previous example, and the information on Z2 and Z4 (Figure 4) relate to each other (Z4 expands Z2).

There are enough white areas (A), so the information is uncluttered; the number of colors is reduced (B), varying from yellow to black to demonstrate levels of severity (seen in the legend, not on the map), a well-known color convention (H) to show levels of severity (I). Also, the number of the map’s subdivisions are reduced in relation to the map in Example 1, decreasing the amount of information to be “read”. However, there is little contrast (F) between orange and red, which could be an issue depending on the quality of the user’s device, as well as to visually impaired readers. Moreover, from yellow to red, there is an intermediate level (orange) to demonstrate severity levels (I) which does not happen from red to black, changing from one color to another dramatically.
Probably, the intention of the black flag is to emphasize the situation’s danger. However, according to the code, there is no possibility to increase the legend if infection risk levels increase.

The legend (G) is clear and identifies the risk level of each region. The meaning and consequences of being classified in each level can only be found on the drop-down menu on Z4. When a level is selected on this menu, the same color appears in a flag, showing similarity (E) and linking the information. The only areas with vibrant colors are the map and the legend, which means the core information is highlighted (C, D).

6 Conclusion

In this study, a framework was developed to analyze colors on maps regarding visual Covid-19 information. The framework, based on the sample collected and on a literature review, has nine categories of analysis and was applied to the sample collected from official governmental websites of the state of Rio Grande do Sul, Brazil.

The first example’s analysis demonstrates possible communication issues as it presents various bright colors in the same section, fails to use contrast to highlight important information, uses similar colors for different purposes, and color shades to represent coded areas are too subtle. The second example was better designed, as it uses colors more consistently, has less colors overall, employs a clear legend, and contains a distinct division of the web page’s information. Another
issue noted is that both examples use similar colors for different information, even though they are from the same governmental organization. This inconsistent use of colors throughout platforms from the same source can be misleading to the general public. In life-threatening situations, where the public is required to act, it is important that the information is consistent and easy to understand.

Given that frameworks were used to analyze this information, it is possible that another researcher could take a different perspective and choose different categories, even with the same sample set. Therefore, for further studies, we expect other researchers to use the framework to validate it. We also suggest extending the framework to consider other features such as typography. Finally, although this framework was created to analyze maps, we believe it also could be used and adapted to analyze other types of visual information.

References


Governo do Estado do Rio Grande do Sul. Mapa de avaliação de risco de cada município e modelo de distanciamento controlado. Available in: https://distanciamentocontrolado.rs.gov.br


Kloh, S. C.; Zimmermann, A. | Framework to analyze the use of colors in informational maps related to Covid-19 data


About the authors

Sara Copetti Klohn
saracopetti@gmail.com
Universidade Federal do Rio Grande do Sul (UFRGS), Brasil

Anelise Zimmermann
anelise.zimmermann@gmail.com
Universidade do Estado de Santa Catarina (UDESC), Brasil

Submission date: 3/5/2021
Approvement date: 2/11/2021