Transparent Risks Revisited: Evidence for a Dark-is-More Bias in Risk Perception

Laura E. Matzen

Sandia National Laboratories

Abstract

Prior research has shown that different representations of uncertainty in data visualizations can lead to more (or less) risk-averse decision making. It is crucial for researchers to develop a better scientific understanding of these effects so that visualizations such as hazard maps can be designed to support viewers in reasoning about risk and probability. This paper presents a follow-up to a prior study that showed that participants underestimated the risk from a wildfire when transparency was used to represent different risk levels. In the present study, we test the hypothesis that the participants' decisions about risk are influenced by the dark-is-more bias. Across three experiments using the same wildfire evacuation task, we found that participants were consistently more likely to evacuate when the probability bands representing the fire risk were darker.

CCS Concepts

• ***Human-centered computing*** → ***Empirical studies in visualization;***

1. Introduction

Although data visualizations are a powerful tool for communicating information about risk and uncertainty [FPS*21, MSD*24, ZC07], prior work has found that they can also produce visual-spatial and cognitive biases in viewers [DFP*18, PCRHS18, PRCR17, PCRT20]. In our prior work [MHTD24], we found that different visualizations of uncertainty led to different patterns of decision making in a simulated wildfire evacuation task. When shown maps depicting fire risk, participants were less likely to evacuate when the risk was encoded via transparency than when it was encoded via hue. The maps included legends that linked each color to the probability of a negative outcome, yet the participants underestimated the risk for perceptually lighter regions. This effect was reduced but not eliminated by the addition of border lines that made the risk bands more distinct.

These findings suggest that participants face two problems when interpreting transparent overlays on hazard maps. First, they may struggle to distinguish light colors when the boundaries between risk levels are not salient [SM13, TLG15]. Second, light colors may bias viewers to underestimate risk. Prior research has identified the *dark-is-more* bias, where viewers infer that darker colors correspond to higher quantities [SGS*18,SRLS20]. However, these studies have focused on differences in response times for *correct* decisions about visualizations with different color mappings. It is not clear whether these biases would impact performance in a task with no time pressure and incentives for careful decision making. Furthermore, the transparent overlays in the wildfire risk maps used darker colors for higher risk levels, consistent with the dark-is-more bias. The participants' poor performance suggests that they were biased by the relative differences in lightness *within* the maps rather than the direction of the color scale. This raises critical questions: Can the dark-is-more bias influence viewers' perception of risk for individual regions within a hazard map? Does this bias override numerical information about risk?

In this paper we present a series of experiments to address these questions. Experiment 1 used two versions of a color scale that varied in lightness, with either light or dark colors associated with higher risk. In Experiments 2 and 3, the colors were mixed so that the risk bands alternated between light and dark. All three experiments found higher evacuation rates when the bands were darker, indicating that the relative darkness of each band had a stronger influence on participants' decisions than the numerical information in the legend. Changing the contrast by using a dark background map (Experiment 3) reduced but did not eliminate this effect. These findings replicate and extend our prior work, demonstrating that transparent map overlays may lead viewers to underestimate risks due to the perception of relative differences in lightness between the colors.

2. Experiment Methods

All experiments were reviewed and approved by the Human Studies Board at Sandia National Laboratories. The data collection took place on Amazon Mechanical Turk. The participants had the Mechanical Turk "master" qualification and were located in the United States. The wildfire evacuation task was the same as in [MHTD24].

DOI: 10.2312/evs.20251077

Participants were shown 800x600 pixel maps with overlays that indicated the probability of each region being in the burn zone of a wildfire, with an X in one region indicating the location of a house. The overlays used the YlOrBr Matplotlib colormap [Hun07], which has an orange hue varying in lightness. Although there are cultural differences in associations between colors and concepts, orange is a fiery color that is also associated with the concept of "hazard" across multiple cultures [OW14].

In our prior work, we consistently found higher evacuation rates for text stimuli (e.g. "there is a 30-40% chance that your house will be in the burn zone") relative to visualizations [MHTD24]. The text stimuli were included in the present experiments to test whether this effect would replicate when darker maps were used.

The trials were self-paced. On each trial, the participants decided whether they would stay or evacuate and clicked a corresponding button to record their decision. A screen showing the outcome of the scenario (safe or burned) was shown for 1.5 seconds, then the next trail began.

Although this task cannot replicate the real-world complexity of making evacuation decisions, we used bonus payments to incentivize careful decision making [CBK*16, MHTD23, MHTD24]. The participants earned 5 cents for each correct decision. If they chose to evacuate, they had to pay 2 cents, representing the real-world costs of evacuating. If they chose to stay and their house burned down, they lost 10 cents, reflecting the higher cost of failing to evacuate from a dangerous situation. Based on this cost/benefit structure, the participants would maximize the expected value of their bonus if they chose to evacuate whenever the probability of their house being in the burn zone was 40% or higher. The participants were not informed of the optimal strategy.

3. Experiment 1: Risk Perception for Darker versus Lighter Colors

Experiment 1 used four types of visualizations, as shown in Figure 1. The background maps used the CARTO "light, no labels" style [Vel22, car]. There were four styles of map overlays, which manipulated the direction of the color scale and the inclusion of salient boundary lines around the probability bands. The stimuli in which darker colors represented higher risk were identical to those used in Experiment 3 of [MHTD24]. For the new stimuli, the color scale was flipped so that lighter colors represented higher risk. As in [MHTD24], the stimuli had three different levels of specificity. The low- medium- and high-specificity stimuli had 3, 6 or 9 risk bands that covered 30%, 15% or 10% ranges of probability.

The stimuli were divided into counterbalanced lists in which every map stimulus was rotated through every visualization condition. The trials were blocked so that the first half of the experiment contained maps with the same color scale and the second half of the experiment contained the maps with the other color scale. The order of the blocks was counterbalanced across participants. Each block contained 54 map stimuli and 13 text stimuli, presented in a different random order for each person. There was a break between the two blocks that informed the participants that they would be switching to the other color scale. They were reminded them to check the legend if they were unsure of what the colors meant.

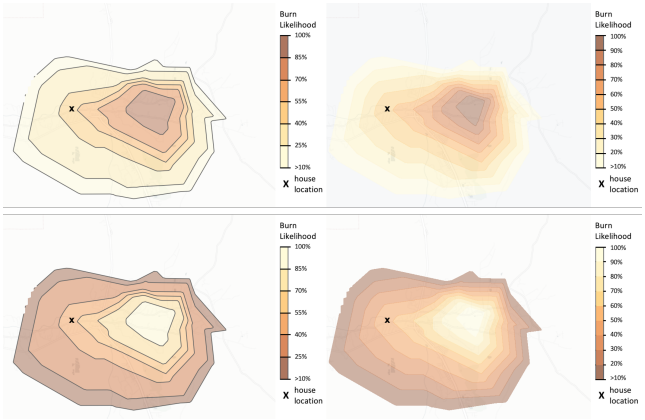

Figure 1: *Examples of a stimulus from Experiment 1 appearing in the four different visualization conditions. The examples with salient border lines show medium-specificity, 6-band stimuli (left), and the examples without salient border lines show high-specificity, 9-band stimuli (right).*

A total of 99 people participated in Experiment 1 and were paid $4.75 for their participation. Three were excluded from the analysis because they ended with a negative bonus, indicating a poor understanding of the task. This left 96 participants who earned an average bonus of $1.73 (range $0.02-$2.32).

3.1. Results

The details of the statistical analyses (including the analyses of the text stimuli) are presented in the Supplemental Materials. As in our prior work, the text condition had the highest overall evacuation rates, with participants evacuating on 66% of trials on average (standard deviation = 10%). When there were no border lines, the average evacuation rates were 57% (SD = 9%) when the low risk bands were lighter and 63% (SD = 10%) when they were darker. The addition of border lines eliminated this difference, with 60% (SD = 9%) and 61% (SD = 11%) evacuation rates for the light and dark conditions, respectively. A 2x2x3 (color scale x borders x specificity level) repeated measures ANOVA was used to compare evacuation rates for the map conditions. Post-hoc test with Bonferroni correction showed that the participants were significantly more likely to evacuate at low risk levels when those bands were dark. This effect was driven by the maps without border lines in the medium- and high-specificity conditions, as shown in Figure 2.

These results are consistent with the findings of [MHTD24] and lend support to the idea that viewers associate lighter colors with lower levels of risk. However, it is still possible that the risk bands were simply harder to distinguish when they were lighter, since the addition of border lines eliminated this effect. To investigate this finding further, we designed Experiment 2, which shuffled the color scale so that the risk bands alternated between light and dark. In this case, the bands were easily distinguished even without border lines. We hypothesized that we would see higher evacuation rates when any given risk band was visualized with a darker color.

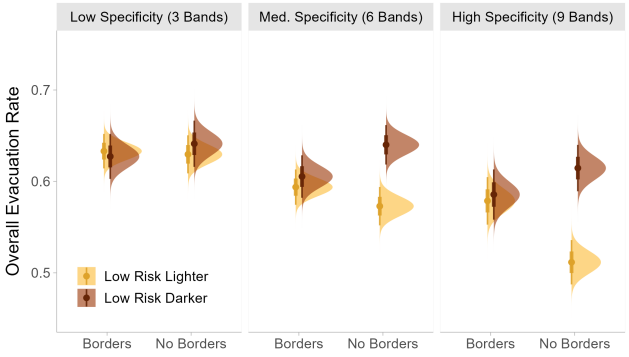

Figure 2: *The evacuation rates for each map condition in Experiment 1. The circles denote the mean and the lines denote the 66% and 95% confidence intervals.*

4. Experiment 2: Mixing Dark and Light Risk Bands

The maps in Experiment 2 used the same colors as before, but the bands were intermixed to alternate between light and dark. We refer to these as the "tiger maps." There were two variants: one where the highest risk band had a light color (Tiger1), and one where the highest risk band had a darker color (Tiger2), as shown in Figure 3 (larger examples are provided in the Supplemental Materials). The low-specificity maps were excluded due to the lack of interesting results for that condition in prior experiments. The participants saw 18 text stimuli and 72 maps with map condition (Tiger1 or Tiger2) as between-subjects variable. Sixty participants completed the task, with 30 seeing each version of the maps. They earned a base payment of $3.75 and an average bonus of $1.11 (range $0.33-$1.72).

4.1. Results

Once again, the average overall evacuation rate was highest for text (67%, SD = 11%), with significantly lower rates for the maps (Tiger1: 61%, SD = 9%; Tiger2 62%, SD = 12%). Between-subjects ANOVAs were used to analyze the effects of band lightness at each level of specificity. The participants were significantly more likely to evacuate when the bands were dark than when they were light. For the medium-specificity maps, the difference in evacuation rates was significant for every risk band *except* for the 55-70% band. For the high-specificity maps, the difference was significant for the 20-30%, 30-40%, 50-60%, 60-70%, and 80-90% bands. These results are shown in Figure 3. The results indicate that the darkness of the colors used to encode each level of risk had a substantial impact on participants' perception of that risk and their subsequent decisions.

5. Experiment 3: Changing the Darkness of the Background

Experiment 3 was designed to assess how much of the effect observed in Experiment 2 was due to the darkness of the bands themselves as opposed to the contrast between the bands and the background. To test this, background maps were changed to the "dark, no labels" style. Examples of the stimuli are shown in Figure 4. Sixty participants completed the task, with 30 seeing each version

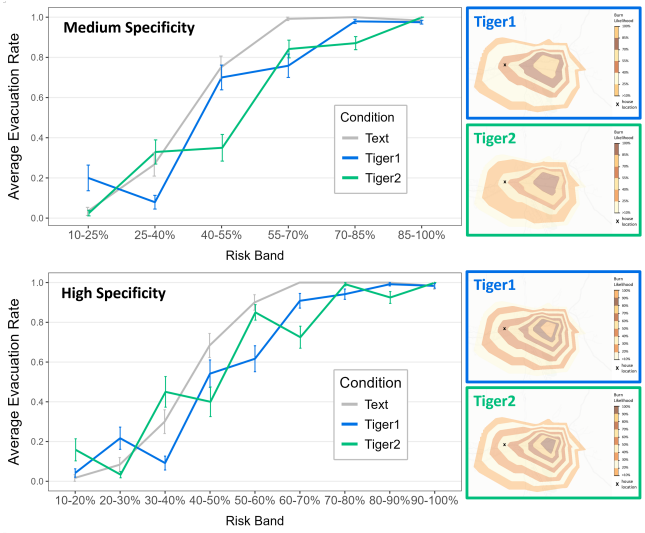

Figure 3: *The evacuation rates for each risk band for the Tiger1 and Tiger2 map conditions in Experiment 2. The error bars show the standard error of the mean.*

5.1. Results

The average overall evacuation rates were highest for text (65%, SD = 11%), but in this case they were not significantly higher than the evacuation rates for the maps (Tiger1: 64%, SD = 8%; Tiger2: 62%, SD = 10%). The statistical analysis showed that the participants were significantly more likely to evacuate when the bands were dark than when they were light. However, when looking at individual risk bands, the difference in evacuation rates between the lighter and darker versions only reached significance for the middle risk bands (the 40-55% risk band for the medium-specificity maps and the 40-50% risk band for the high-specificity maps). These results are shown in Figure 4.

The combined data from Experiments 2 and 3 were analyzed with a three-way ANOVA testing the effects of band color, background color, and specificity. This analysis showed significant main effects of band lightness and specificity, plus a significant interaction between band lightness and background color. Post-hoc tests showed that for darker bands, the evacuation rates did not differ significantly across the dark and light background maps. However, the participants were significantly more likely to evacuate when they saw light bands on a dark background compared to when they saw light bands on a light background. This was true for both medium- and high-specificity risk maps. These results are shown in Figure 5

5.2. Discussion

The results of these experiments provide new evidence that the relative lightness or darkness of the colors on a hazard map can influence participants' perceptions of risk and their subsequent de-

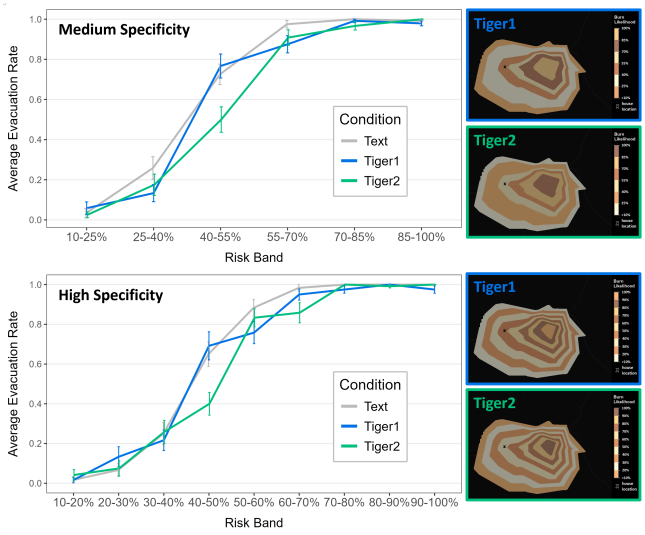

Figure 4: *The evacuation rates for each risk band for the Tiger1 and Tiger2 conditions in Experiment 3. The error bars show the standard error of the mean.*

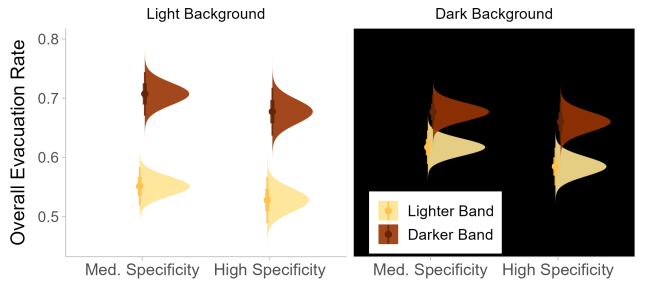

Figure 5: *The overall evacuation rates for each map condition in Experiments 2 and 3.*

cisions. All three experiments found the same pattern of performance, with participants making more risk-averse decisions in response to darker risk bands. These findings are consistent with the dark-is-more bias, where people infer that darker colors map to larger quantities [Cuf73, McG89, SGS*18, SRLS20]. Our findings indicate that this principle extends to risk perception and that this bias can lead to different decisions even when the numerical risk information is identical. Importantly, the biases in risk perception appear to be driven by the differences in lightness between risk bands rather than the direction of the color scale. This has important implications for transparent map overlays, since variations in transparency may be perceived as variations in lightness [Bec85, SA02, WPH17].

Changing the color of the background map from light to dark mitigated, but did not eliminate, the dark-is-more bias. This is consistent with prior work that found that the dark-is-more bias persists but is reduced when the background is dark [McG89]. In our case, this change increased the contrast between the light bands and the background, but it also made the transparent overlays appear darker overall. Notably, this also eliminated the difference between the text and visualization stimuli in our study (additional discussion of the text stimuli can be found in the Supplemental Materials). It is not clear how much these effects were driven by the change in overall lightness versus the change in relative lightness and/or contrast. Contrast is an important variable that is known to impact perception of maps [Bre92, CRB05, ŠB*15]. Our findings suggest that the darkness of the risk bands played a larger role than their contrast with the background of the hazard maps, but there may be other combinations of visual features where contrast would override lightness. More research in this area is needed to better understand how these visual cues influence risk perception.

5.3. Limitations

An important limitation of this research is that the data were collected online, so we had no control over the brightness and contrast settings of the participants' displays. Future research using more tightly-controlled display settings would be beneficial. Another limitation is our use of the YlOrBr colormap, which varies in lightness but also in hue. Lightness and hue may interact to influence risk perception in ways that are not apparent in the present studies. In future work, we plan to use color schemes that vary in *only* lightness or *only* hue.

Another major limitation of this work is that the risks are not real, so the participants' performance may not reflect real-life decisions about risk. However, our experiments indicate that the effect of lightness on risk perception is persistent and powerful. The experimental design revealed that the *same individuals* made more risk-averse decisions when their house was in a darker band compared to when it was in a lighter band, even in cases where the darker risk band had a numerically lower probability of being in the burn zone. These patterns indicate that the bias toward interpreting darker colors as representing greater risk is pervasive, and that it is likely to influence real-world decisions as well.

5.4. Conclusions

Hue, lightness, and transparency are common visual variables in cartography in general and for hazard maps in particular. Widely-used guidelines recommend using hue to encode qualitative differences and lightness to encode quantitative differences [Bre94]. Transparency is extremely common in map overlays because it does not obscure the background map and it can be an effective cue for communicating information about uncertainty [MRO*12]. However, our findings suggest that lightness and transparency (which may be perceived as lightness) should be used carefully when they are used to provide information about uncertain risks, as they may lead to biases in risk perception.

6. Acknowledgments

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

References

[Bec85] BECK J.: Perception of transparency in man and machine. *Computer Vision, Graphics, and Image Processing 31*, 2 (1985), 127–138. doi:10.1016/0734-189X(85)90169-0. 4

[Bre92] BREWER C. A.: Review of colour terms and simultaneous contrast research for cartography. *Cartographica: The International Journal for Geographic Information and Geovisualization 29*, 3-4 (1992), 20–30. doi:10.3138/80ml-3k54-0204-6172. 4

[Bre94] BREWER C. A.: Color use guidelines for mapping and visualization. *Modern Cartography Series 1994*, 123-147 (1994), 7. doi:10.1016/b978-0-08-042415-6.50014-4. 4

[car] Carto. https://carto.com/. Accessed: 2022-02-21. 2

[CBK*16] CHEONG L., BLEISCH S., KEALY A., TOLHURST K., WILKENING T., DUCKHAM M.: Evaluating the impact of visualization of wildfire hazard upon decision-making under uncertainty. *International Journal of Geographical Information Science 30*, 7 (2016), 1377–1404. doi:10.1080/13658816.2015.1131829. 2

[CRB05] CHESNEAU E., RUAS A., BONIN O.: Colour contrasts analysis for a better legibility of graphic signs on risk maps. In *Proc. 22th International Cartographic Conference (ICC)* (2005). 4

[Cuf73] CUFF D. J.: Colour on temperature maps. *The Cartographic Journal 10*, 1 (1973), 17–21. doi:10.1179/000870473787339347. 4

[DFP*18] DIMARA E., FRANCONERI S., PLAISANT C., BEZERIANOS A., DRAGICEVIC P.: A task-based taxonomy of cognitive biases for information visualization. *IEEE Transactions on Visualization and Computer Graphics 26*, 2 (2018), 1413–1432. doi:10.1109/tvcg.2018.2872577. 1

[FPS*21] FRANCONERI S., PADILLA L., SHAH P., ZACKS J., HULLMAN J.: The science of visual data communication: What works? *Psychological Science in the Public Interest 22*, 3 (2021), 110–161. doi:10.1177/15291006221099494. 1

[Hun07] HUNTER J. D.: Matplotlib: A 2d graphics environment. *Computing in Science & Engineering 9*, 3 (2007), 90–95. doi:10.1109/MCSE.2007.55. 2

[McG89] MCGRANAGHAN M.: Ordering choropleth map symbols: The effect of background. *The American Cartographer 16*, 4 (1989), 279–285. doi:10.1559/152304089783813918. 4

[MHTD23] MATZEN L. E., HOWELL B. C., TRUMBO M. C. S., DIVIS K. M.: Numerical and visual representations of uncertainty lead to different patterns of decision making. *IEEE Computer Graphics and Applications 43*, 5 (2023), 72–82. doi:10.1109/MCG.2023.3299875. 2

[MHTD24] MATZEN L. E., HOWELL B. C., TUFT M., DIVIS K. M.: Transparent risks: The impact of the specificity and visual encoding of uncertainty on decision making. In *Computer Graphics Forum* (2024), vol. 43, Wiley Online Library, p. e15094. doi:10.1111/cgf.15094. 1, 2

[MRO*12] MACEACHREN A. M., ROTH R. E., O'BRIEN J., LI B., SWINGLEY D., GAHEGAN M.: Visual semiotics & uncertainty visualization: An empirical study. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (2012), 2496–2505. doi:10.1109/tvcg.2012.279. 4

[MSD*24] MATZEN L. E., STITES M. C., DIVIS K. M., BENDECK A. E., STASKO J. T., PADILLA L. M.: Effects of forecast order, cost, and risk on decision making with multiple forecast visualizations. In *2024 IEEE Workshop on Uncertainty Visualization: Applications, Techniques, Software, and Decision Frameworks* (2024), IEEE, pp. 28–37. doi:10.1109/uncertaintyvisualization63963.2024.00008. 1

[OW14] OR C. K., WANG H. H.: Color–concept associations: A cross-occupational and-cultural study and comparison. *Color Research & Application 39*, 6 (2014), 630–635. doi:10.1002/col.21832. 2

[PCRHS18] PADILLA L. M., CREEM-REGEHR S. H., HEGARTY M., STEFANUCCI J. K.: Decision making with visualizations: a cognitive framework across disciplines. *Cognitive Research: Principles and Implications 3*, 1 (2018), 1–25. doi:10.1186/s41235-018-0120-9. 1

[PCRT20] PADILLA L. M., CREEM-REGEHR S. H., THOMPSON W.: The powerful influence of marks: Visual and knowledge-driven processing in hurricane track displays. *Journal of Experimental Psychology: Applied 26*, 1 (2020), 1. doi:10.31234/osf.io/5tg9y_v1. 1

[PRCR17] PADILLA L. M., RUGINSKI I. T., CREEM-REGEHR S. H.: Effects of ensemble and summary displays on interpretations of geospatial uncertainty data. *Cognitive Research: Principles and Implications 2* (2017), 1–16. doi:10.1186/s41235-017-0076-1. 1

[SA02] SINGH M., ANDERSON B. L.: Toward a perceptual theory of transparency. *Psychological Review 109*, 3 (2002), 492. doi:10.1037//0033-295x.109.3.492. 4

[ŠB*15] ŠTĚRBA Z., BLÁHA J. D., ET AL.: The influence of colour on the perception of cartographic visualizations. *Color and Image* (2015), 533–38. 4

[SGS*18] SCHLOSS K. B., GRAMAZIO C. C., SILVERMAN A. T., PARKER M. L., WANG A. S.: Mapping color to meaning in colormap data visualizations. *IEEE Transactions on Visualization and Computer Graphics 25*, 1 (2018), 810–819. doi:10.1109/tvcg.2018.2865147. 1, 4

[SM13] SEVERTSON D. J., MYERS J. D.: The influence of uncertain map features on risk beliefs and perceived ambiguity for maps of modeled cancer risk from air pollution. *Risk Analysis 33*, 5 (2013), 818–837. doi:10.1111/j.1539-6924.2012.01893.x. 1

[SRLS20] SIBREL S. C., RATHORE R., LESSARD L., SCHLOSS K. B.: The relation between color and spatial structure for interpreting colormap data visualizations. *Journal of vision 20*, 12 (2020), 7–7. doi:10.1167/jov.20.12.7. 1, 4

[TLG15] THOMPSON M., LINDSAY J., GAILLARD J.: The influence of probabilistic volcanic hazard map properties on hazard communication. *Journal of Applied Volcanology 4*, 6 (2015). doi:10.1186/s13617-015-0023-0. 1

[Vel22] VELARDE V.: Cartodb basemap styles. https://github.com/CartoDB/basemap-styles, 2022. 2

[WPH17] WANG C., POULIOT J., HUBERT F.: How users perceive transparency in the 3d visualization of cadastre: testing its usability in an online questionnaire. *GeoInformatica 21* (2017), 599–618. doi:10.1007/s10707-016-0281-y. 4

[ZC07] ZUK T., CARPENDALE S.: Visualization of uncertainty and reasoning. In *Smart Graphics* (2007), Butz A., Fisher B., Krüger A., Olivier P., Owada S., (Eds.), Springer, pp. 164–177. doi:10.1007/978-3-540-73214-3_15. 1