DISSERTATION

SHADES OF RISK: A MIXED-METHODS APPROACH TO DESIGNING AND TESTING A NEW HURRICANE MAP GRAPHIC

Submitted by

Zoey Rosen

Department of Journalism and Media Communication

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2023

Doctoral Committee:

Advisor: Marilee Long

Katie Abrams
Gayathri Sivakumar
Sheryl Magzamen
David Most

ABSTRACT

SHADES OF RISK: A MIXED-METHODS APPROACH TO DESIGNING AND TESTING A NEW HURRICANE MAP GRAPHIC

Map graphics are a popular tool for hazard risk communication, layered with numerical, verbal, and visual information to describe an uncertain threat. In the hurricane context, graphics are used to communicate the probability of different threats over a forecasting period. While hurricane graphics have been studied in the past, they have not been analyzed from the design phase through to the intended audience. Additionally, hurricane graphics have not been designed with colorblind-friendly accessibility in mind. This dissertation presents the results of a three-phase, mixed methods study: (a) graphic development, (b) testing with expert user groups, and (c) testing with a public sample. In the development phase (a), I used the best practices for using probability language, color schemes, and localization into map graphics from literature in forecasting, communication, universal design, and emergency management. Additionally, I held informal interviews with professionals from the National Hurricane Center to develop the prototype with their recommendations for the design. In the first testing phase b, I interviewed 19 expert users (emergency managers and meteorologists) from Florida and Louisiana about their preferences for and feedback on the design elements of a new hurricane graphic, as well as if there were individual characteristics that influenced how accurate they were in interpreting wind exceedance data, such as risk perception, confidence, experience, spatial cognition, and numeracy levels. In phase c, I tested the wind exceedance graphic prototypes using a public sample (n = 624) from Louisiana and Florida to gather data on the accuracy of their

interpretations for the graphic, again measuring confidence, experience, spatial cognition, and numeracy levels, as well as their design preferences and risk perceptions.

The results of the two testing phases (b and c) center around how accurate experts and the public were with interpreting the graphic, as well as if there were other factors that influenced this accuracy, such as spatial cognition or numeracy. Additionally, the results describe both groups' design preferences, risk perceptions of the color schemes and overlays, and how experts think about vulnerability when using the graphic. In both groups, numeracy and spatial cognition were found to predict accuracy of interpretation for a wind exceedance graphic prototype. Likewise, both confidence and experience were found to have a positive relationship with accuracy. Regarding the design choices, both experts and the public preferred a yellow-to-red scheme, though experts thought the yellow-to-red scheme presented the hazard as riskier and the public thought the reds-only was riskier. Adding overlays to the graphic, such as interstates or city landmarks, helped the participants to orient themselves on the map. Experts and the public preferred that there were overlays added to the graphic and scored this version of the graphic as risker than a version without any overlays. The addition of the overlays prompted expert users to think more about the risk and vulnerability of the people in those areas on the map. Vulnerability was conceptualized from both a physical and social standpoint by the experts and applied to how they would use the wind exceedance graphic in a briefing to communicate to their community partners. Overall, this research provides a model for how hazard risk map graphics can be studied from design through implementation. Additionally, I captured how experts think about vulnerability in their communities when shown a forecast map graphic. The conclusion of this dissertation also provides practical recommendations for experts who want to apply the universal design aspects into new hurricane graphics.

ACKNOWLEDGMENTS

This dissertation and the last six years of graduate school would not have been possible without the help and support of some incredible people. Thank you to my advisor, Dr. Marilee Long. I truly cannot put into words how much I am inspired by you and value your mentorship. I am so grateful for every meeting in your office where we got to talk through theory, methods, and grammar rules, as well as telling stories when we needed to change topics.

Thank you to my committee, Dr. Katie Abrams, Dr. Gayathri Sivakumar, Dr. Sheryl Magzamen, Dr. David Most, and Andrea Schumacher. Each class and research meeting helped to make me a better scholar. I appreciate all the time you took to read every draft of this behemoth of a document and am honored that I can list your names alongside mine on the title page.

Thank you to my parents (Beth and Connie) and sister (Katy) for their support, answering the phone each time I wanted to procrastinate, and pictures of Remi and Watson for motivation. Your endless support to get me through graduate school means the world to me. I learned to ask questions from a young age at home and look how far it has taken me. Thank you to my family, in alphabetical order, the Beckers, Fjelsted-Cohns, Kanes, Rosens, and Solberg-Quinns for your support in this (and every!) hurricane season.

Thank you to my wonderful friends for every brunch and trip to a brewery that made this PhD possible. Thank you especially to Aaron Alexander and Channing Bice for listening to me talk through ideas, read my paragraphs aloud, and brainstorming ideas for figures. As for the rest of my amazing friends to which there are so many to list: you know who you are and that I love and appreciate you, I am just very tired of writing.

TABLE OF CONTENTS

ABSTRACT II

ACKNOWLEDGMENTS IV

LIST OF TABLES VIII

LIST OF FIGURES IX

LIST OF ACRONYMS X

CHAPTER 1. INTRODUCTION 1

1.1 Communication Elements in Graphical Forecast Products 5
1.2 Study Purpose 8

CHAPTER 2. LITERATURE REVIEW 9

2.1 Risk Perception 9
 2.1.1 Experience and Risk Perception 14
2.2 Universal Design in Hazard Visualizations 17
 2.2.1 Probability Language 19
 2.2.2 Color Choice and Texture 22
 2.2.3 Localization 27
2.3 Vulnerability and Decision-Making 32

CHAPTER 3. METHODS 37

3.1 Graphic Development 37
 3.1.1 Informal Interview with NHC Liaisons 41
3.2 Testing of the Prototype with Expert Users 42
3.3 Public Testing of the Graphic Prototypes 46
3.4 Measurement of Variables of Interest 48
 3.4.1 Numeracy 48
 3.4.2 Spatial Cognition 49
 3.4.3 Experience with Hurricanes 51
 3.4.4 Accuracy of Interpretation 54
 3.4.5 Confidence in Interpretation 55
 3.4.6 Risk Perceptions 56
 3.4.7 Metacognitions of Vulnerability 59
3.5 Expert Test Procedures (Interviews) 59
3.6 Public Test Procedures (Survey) 61
3.7 Analytic Approach 62
3.8 Reflections from the Pilot Tests of the Public Test and Expert Test 63
3.9 Expert User Sample 66
 3.9.1 Reliability and Normality of Expert Quantitative Data 67

3.10 Public Sample 67
3.10.1 Reliability and Normality in Public Quantitative Data 68

CHAPTER 4. RESULTS 71

4.1 Update to Methodological Approach 71
4.2 Findings Related to Hurricane Experience 71
4.2.1 Hurricane Experience and Risk Perceptions in Expert Users 71
4.2.2 Hurricane Experience and Risk Perceptions in the General Public 75
4.2.3 Hurricane Experience and Accuracy of Interpretation in Experts and the Public 76
4.3 Findings Related to Confidence in Using Wind Exceedance Graphic 78
4.3.1 Relationship between Numeracy and Confidence in Wind Exceedance Interpretation for Experts and the Public 79
4.4 Findings Related to Color Scheme for the Wind Exceedance Graphic 79
4.4.1 Expert User Color Preferences 79
4.4.2 General Public Color Preferences 82
4.4.3 Effect of Color Scheme on Perceived Riskiness of Graphic 82
4.5 Findings Related to Overlays in the Wind Exceedance Graphic 86
4.5.1 Expert User Overlay Preferences 86
4.5.2 Public Sample Overlay Preferences 89
4.5.3 Perceived Riskiness of Wind Exceedance Graphic with Overlays 90
4.6 Findings Related to Spatial Cognition 92
4.7 Expert Metacognitions of Vulnerability in Hurricane-prone Areas 94
4.7.1 Physical Vulnerability 94
4.7.2 Social Vulnerability 97

CHAPTER 5. DISCUSSION 100

5.1 Accuracy, Numeracy, Spatial Cognition, and Confidence on Graphic Interpretation 102
5.2 Experience and Risk Perception 105
5.3 Perceptions and Impressions of Universal Design Elements in the Wind Exceedance Graphic 108
5.3.1 Color Preference and Perceived Risk in Experts and the Public 108
5.3.2 Overlay Preference and Perceived Risk in Experts and the Public 111
5.4 Vulnerability and Communicating Effectively to Diverse Audiences 113
5.5 Limitations 118

CHAPTER 6. CONCLUSION 122

6.1 Future Research 124
6.2 Practical Recommendations 130

REFERENCES 133

APPENDIX 144

Subjective Numeracy Scale (SNS) (Fagerlin et al., 2007) 144
Santa Barbara Sense of Direction (SBSOD) Scale (Hegarty et al., 2002) 145
Risk Perception Items (Wilson et al., 2019) 146

Hurricane Experience Score (HES) (Ehrlich et al., 2010) 147
Interview Recruitment Email 148
Interview Oral Informed Consent Script.. 149
Interview Guide 151
Survey Informed Consent 153
Other Demographic Data from Public Survey 154
Prototype Graphics.. 156
Accuracy Measure Explanatory Figure 159

LIST OF TABLES

Table 1. Reliability (alpha) and Normality (W) Statistics for Expert Sample Quantitative Data. 67
Table 2. Reliability Statistics for Public Quantitative Data. 68
Table 3. Independent samples t-test values for scales in public survey data. 69
Table 4. Descriptive statistics for numeracy [6-point] and confidence [5-point] measures. 79
Table 5. Descriptive statistics for riskiness [5-poimt scale] of color schemes, by preference group in public sample. 85
Table 6. Wilcoxon Signed Ranks test for perceived riskiness [5-point scale] of overlays in public sample. 91
Table 7. Descriptive statistics for numeracy [6-point scale] and spatial cognition [7-point scale] in both samples. 92

LIST OF FIGURES

Figure 1. Explanatory figure depicting cumulative probability of wind exceedance with max surface wind ... 2
Figure 2. Example of current wind speed probability time-of-arrival hurricane forecast product for Hurricane Ida in 2021 ... 6
Figure 3. Proposed color schemes for the wind exceedance product ... 24
Figure 4. Examples of color and texture schemes for the wind exceedance graphic prototypes . 25
Figure 5. Cone of Uncertainty sample graphic (National Hurricane Center & Central Pacific Hurricane Center, 2022)... 26
Figure 6. Close-up of landmarks on wind exceedance graphic ... 28
Figure 7. Overlapping histogram of number of years participants report living in their state ... 70
Figure 8. Histogram of hurricane experience score in public sample... 76
Figure 9. Boxplots of accuracy scores for the wind exceedance graphic for public and expert samples... 78
Figure 10. Stacked bar chart of color scheme preference by group... 82
Figure 11. Boxplots comparing differences in perceived riskiness of the texture scheme, by color scheme preference group ... 86
Figure 12. Tweet from @NWSOmaha about how often they get questions about unique cities . 87
Figure 13. Stacked bar chart of overlay preference, showing city landmarks was largely preferred by the public and the experts were about equal in preference between interstates/highways and city landmarks. ... 90
Figure 14. Visual diagram of how accuracy score was determined from participant responses 159

LIST OF ACRONYMS

- American Meteorological Society (AMS)
- Cooperative Institute for Research in the Atmosphere (CIRA)
- Emergency Managers (EMs)
- Federal Emergency Management Agency (FEMA)
- Geographic information systems (GIS)
- Hazardous Weather Testbed (HWT)
- Hurricane Local Statement (HLS)
- Hurricane and Ocean Testbed (HOT)
- Hurricane Experience Score (HES)
- Intergovernmental Panel on Climate Change (IPCC)
- Joint Technology Transfer Initiative (JTTI)
- National Hurricane Center (NHC)
- National Oceanic and Atmospheric Administration (NOAA)
- National Weather Service (NWS)
- Risk information seeking and processing (RISP)
- Santa Barbara Sense of Direction (SBSOD)
- Subjective Numeracy Scale (SNS)
- Spatial Thinking Ability Test (STAT)
- Social Vulnerability Index (SVI)
- Universal design (UD)
- Wind speed Probability (WSP)
- Wind speed Probability-based Tropical Cyclone Message (WTCM)

CHAPTER 1. INTRODUCTION

Hurricane forecasting is more than data assimilation and modeling; meteorologists from the National Hurricane Center (NHC) not only need to produce high-quality probability information, but they need to communicate the uncertainties within their forecasts to important stakeholders, like emergency managers (EMs) and meteorologists in affected forecasting offices and broadcast news stations, as well as directly to the public. In turn, these EMs and meteorologists become science communicators for their audiences during hurricanes. The challenge of weather prediction in addition to the challenge of predicting human behavior sets the stage for just how complicated, yet critical it is to communicate severe weather forecasts effectively. Thus, it is important to understand how these groups interpret meteorological data products, since how they understand this information influences how they explain their decisions to their audiences.

The mission of the Cooperative Institute for Research in the Atmosphere (CIRA) is to reduce the challenges associated with creating accurate forecasts. CIRA does this by conducting multi-disciplinary research projects that create new, effective products and models. In August 2021, CIRA started a new Joint Technology Transfer Initiative (JTTI)-funded project in collaboration with the National Weather Service (NWS) Miami/South Florida weather forecasting office and the NHC. One goal of this project is to develop a new wind exceedance graphic from the WTCM (Wind speed Probability-based Tropical Cyclone Message) model. A wind exceedance graphic should help meteorologists and EMs make better decisions before and during a hurricane.

The current windspeed probability (WSP) graphic depicts the probability (likelihood, expressed as a percentage) that sustained (1-minute average) winds meeting or exceeding specific thresholds will occur at particular locations over specific intervals of time (National Hurricane Center, n.d.). The WSP graphic represents a probability of exceedance, while the new wind exceedance values in the WTCM in this study represent actual windspeed values (mph or knots) that correspond to pre-specified probabilities of exceedance. The new wind exceedance graphic is calculated from the same 1,000 wind model ensemble members (models) used to calculate the WSP. Wind exceedance describes windspeed values along the cumulative exceedance probability curve, corresponding to certain percentile values. For example, if we are interested in finding the predicted surface windspeed value for which there is a 1 in 10 chance of being exceeded, we would look at point P_{10} (Figure 1). There would be a 10% chance that the windspeed observations would be faster than the maximum predicted surface wind at that point.

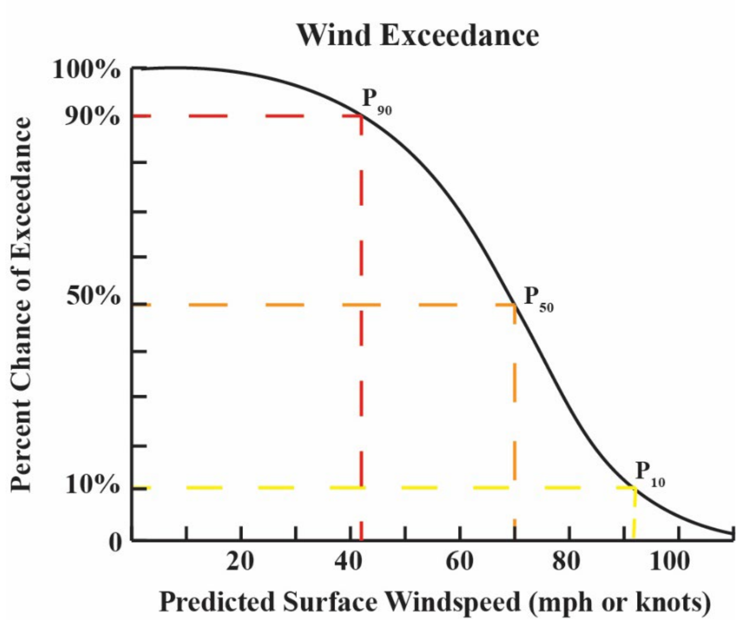


Figure 1. Explanatory figure depicting cumulative probability of wind exceedance with maximum surface wind

One can think of exceedance descriptively like the upper bounds of a confidence interval—there would be a percentage chance that the forecasted wind would be higher than an average predicted range. Wind exceedance is often depicted in different probabilistic ranges, e.g., 10% illustrating the worst-case scenario, 50% illustrating the average scenario, and 90% illustrating the most likely scenario. A graphic showing the 10% depiction of wind exceedance shows the strongest winds likely for the most reasonable, worst-case scenario[1]. Communicating a high-consequence, low-probability risk is challenging, especially for natural disasters; though a 10% probability scenario is less likely to be experienced, the consequences are grave (Bostrom et al., 2008). For example, major earthquakes are highly destructive though they are not typically experienced annually (Bostrom et al., 2008). An example relevant to this study is that the U.S. Navy decides whether to move their ships based on the probability that wind speed will exceed 50 knots per hour (57.54 miles per hour) because of the damage that occurs to ships at that strength (Hogsett, personal communication). While wind exceedance data has been available in past wind speed probability models, EMs have said they would like a new graphic that solely displays this information, especially for the forecasted worst-case scenarios.

Graphic design is interdisciplinary in nature, as it combines scientific data with data visualization techniques that support information processing and effective communication. One area of social science that is particularly salient to the graphic design space is the discipline of mass communication, which studies the dissemination of information in various formats across mediated channels to large populations. To support the JTTI project, a team of researchers from CIRA and Colorado State University's Department of Journalism and Media Communication

[1] "Reasonable-worst case scenario" is a term commonly used by meteorologists when talking to EMs to describe a 10% likelihood.

have developed a new, effective product to show wind exceedance values from an updated model. My research goals concern the graphic and its design, whereas CIRA team members' goals center around the mathematical components of the updated model.

In order to create a graphic that is interpreted accurately and is useful for decision-making, I have (a) designed a wind exceedance forecast graphic based on the science communication and visual processing literature, (b) tested the graphic prototypes with expert user groups (meteorologists and EMs), (c) tested the graphic with the public, and made recommendations for changes to the graphic based on their feedback. Research is needed to determine if both experts (specialized groups) and novices (general public) can interpret the graphic accurately, as expertise influences people's risk assessments and their confidence in these assessments (Roth, 2009). If this one graphic is meant to communicate widely about the wind risk from hurricanes, then it is critical we test it using a sample of users who do not have specialized training or experience with reading forecast maps. In a study on floodplain mapping, Roth (2009) found that when less experienced map users were shown a map displaying uncertainty about flooding, they underestimated the flooding risk for a site on the map compared to more expert users. Not only were novice assessments less accurate, but novice users were also less confident in their assessments (Roth, 2009).

Ideally, assessment accuracy and confidence should both be high; a mismatch could have tragic consequences (Roth, 2009). As meteorologists and EMs routinely interpret forecast uncertainty in maps, they likely have more confidence in their ability to accurately interpret probability information from these graphics. Testing how well the public interprets the wind exceedance graphic would provide more insight into this group's accuracy and confidence in interpreting uncertainty from a graphic. Additionally, a public sample would provide more information about the risk perceptions for wind exceedance. The feedback gained from studying how meteorologists, EMs, and a public sample interpret the wind exceedance product will provide insight into what changes should be made to the graphic before it is fully adopted into operations.

1.1 Communication Elements in Graphical Forecast Products

There are many communication strategies meteorologists use in hurricane briefings, which are hybrid (in-person and virtual) meetings where meteorologists disseminate new hurricane forecasts to stakeholders like EMs and other community partners. Meteorologists prepare a mixture of graphics, text, and maps to help explain the tropical storm conditions. After the briefing, these items may be shared with the public in broadcast weather reports, on the NHC website, through local weather forecasting office social media channels, and more. Because they are spatially distributed, environmental hazards are often represented using maps (Bostrom et al., 2008). Maps also assist in making environmental threats appear more local. In a study of coastal sea-level rise, Retchless (2014) found that overlaying neighborhood and city block grids onto a map of the impacted shoreline helped to make climate change more local and more of a personal threat to residents (Retchless, 2014). Localization is a powerful communication strategy because it personalizes the risks for each person who uses the hazard map (Monmonier, 2008).

As maps are tools that support risk assessment and decision-making, maps depicting environmental hazards can communicate risk and uncertainty effectively (Roth, 2009). Forecasters combine several methods for communicating uncertainty into a single map graphic. For instance, a typical graphic contains (a) a map with geographic scaling to show the spatial distribution of the forecasted hazard's potential effects, (b) a probabilistic numerical range in the title or legend to show the percent chance of a hazard's impacts, and (c) a color scheme to display the differences between categories of the likelihood of an expected hazard. An example of a current wind speed probability and time-of-arrival forecast product is shown in Figure 2.

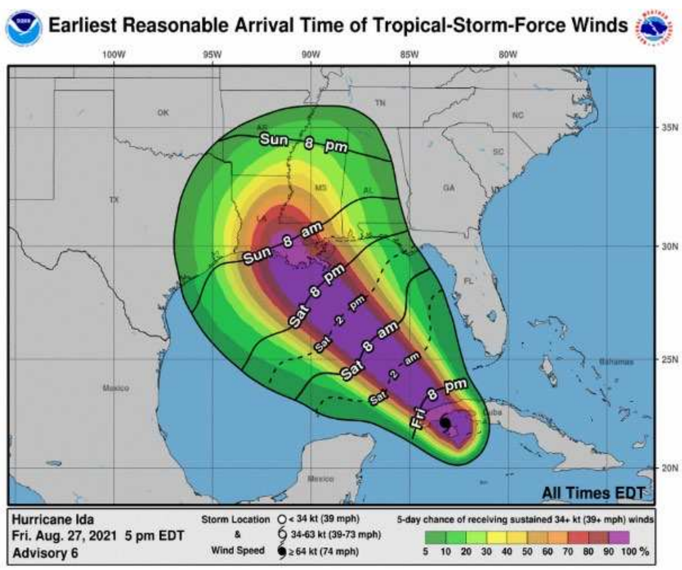


Figure 2. Example of current wind speed probability time-of-arrival hurricane forecast product for Hurricane Ida in 2021

Forecast uncertainty has historically been communicated in a probabilistic format, using a numerical range of hazard severity or timing, for example, so that audience members can assess their own risks from the impending threat (De Elía & Laprise, 2005; Nadav-Greenberg & Joslyn, 2008). However, a probabilistic format is only one quantitative technique for communicating uncertainty, often producing mixed results. For example, evidence suggests that people misinterpret probability in forecasts, interpreting the percent chance (e.g., 20% of rain) as forecaster confidence instead of hazard likelihood (Joslyn et al., 2009; National Academy of Sciences, 2006). In response, forecasters began using just verbal expressions of probability in written forecasts, such as "slight" chance. However, the interpretations for these labels varied and led to more confusion, according to Ripberger et al.'s (2022) review of studies that empirically examined risk messages that used probability information. Ripberger et al. (2022) gave suggestions for how to communicate risk more effectively based on the syntheses of the studies included in their review. For example, two studies included in Ripberger et al. (2022) found that more accurate interpretations of hazard likelihood occur when forecasters include numbers alongside descriptive words (Lenhardt et al., 2020; Rosen et al., 2021). Presented clearly, including numerical probability information generally improves decision quality, which is important for weather threats (Morss et al., 2010; Ripberger et al., 2022).

Like with the presentation of numerical information, the role of color in risk perception is also widely studied (Bostrom et al., 2008). Different color schemes have been tested to see how they convey risk, with the colors red and yellow communicating higher risk than other colors, such as green (Bostrom et al., 2008; Cheong et al., 2020). Color has social dimensions that influence their perception. For example, the finding that red and yellow communicate risk makes sense, as these colors mirror the messages communicated from traffic stoplights where driving through a red light or a yellow light is dangerous, while driving through a green light is safer. While the traffic light color scheme works for some maps, it does not work for every hazard map. Research has shown that using colors that make sense for the phenomenon being presented is important (Bostrom et al., 2008; Cheong et al., 2020). Using blue, for example, would not convey the same risk for wind speed as it would for flooding, because water is commonly associated with the color blue.

Furthermore, just as the color itself is important, so too is its hue and saturation. Varying saturation from low to high (e.g., from light to dark red) in a color scheme can communicate an increasing certainty of a forecasted hazard, while varying hues (e.g., from yellow to red) can communicate that there are different categories or values for each color shown (Maceachren et al., 2012).

Varying the saturation of a single color is also labeled as a univariate color scheme. Varying hues can be labeled as a bivariate color scheme, since the scheme consists of two colors at the lowest and highest categories, with variations of these colors forming the middle values. Comparing univariate to bivariate schemes has been analyzed before in risk maps, though not extensively (e.g., Aerts et al., 2003; Roth, 2009). In a study by Aerts et al. (2003), univariate schemes were generally preferred by both expert and novice users of hazard risk maps. However, experts did have an affinity towards the bivariate color schemes (Aerts et al., 2003). Whether there is one color or many colors in a hazard map, color variations in data visualizations have been shown to alter people's risk perceptions (Bostrom et al., 2008; Cheong et al., 2020). Understanding more about which schemes are effective or preferred by different user groups can aid in creating better graphics in future products (Bostrom et al., 2008).

1.2 Study Purpose

This is the first time that a hurricane map graphic has been developed using universal design principles, and only the second time that a hurricane map graphic was tested with experts and the public in a mixed-methods, social-scientific approach. While this study uses wind exceedance as the prototype graphic, the design can be applied to different hurricane hazards as well. Additionally, this dissertation adds to the knowledge of how experts think about vulnerability in their communities when they are shown a hazard risk map. Therefore, this dissertation is unique in two ways: this study provides a guide for future researchers on how to design an interdisciplinary, mixed-methods project for testing new forecast map graphics with their intended audiences, and this study begins to capture how the presentation of the information in a map affects decision-making by EMs and meteorologists.

CHAPTER 2. LITERATURE REVIEW

This chapter covers scholarship on universal design (UD), including concepts such as probability, color, and localization. These concepts are connected with individual factors (e.g., numeracy and spatial cognition) that could influence risk perceptions, accuracy, confidence, and metacognitions for decision-making. These individual factors often arise in the risk communication scholarship and are elaborated on using examples from the domains of emergency management, finance, geography, health, meteorology, and/or risk communication.

2.1 Risk Perception

Risk is often defined objectively as the probability of an event times the magnitude of adverse consequences (Bostrom et al., 2008; Renn, 2011). Risk perception, in contrast, describes a person's subjective risk assessment, which refers to their evaluation of the likelihood of a hazard with the potential severity of harm (Griffin et al., 2008). Risk perceptions influence how communities assess risk, manage threats, use decision support services[2], and communicate about hazards (Siegrist & Árvai, 2020). There are different approaches to studying risk perception from the cognition scholarship. Three of the most popular approaches include systematic processing, heuristic processing, and psychometrics. While this study focuses on the systematic processing approach, all three approaches are described below.

To begin in the systematic processing literature, risk perception, also referred to sometimes as risk judgment (Ahn & Noh, 2020), describes one's assessment of how likely a

[2] Decision support services "provide forecast advice and interpretative services the NWS provides to help core partners, such as emergency personnel and public safety officials, make decisions when weather, water and climate impacts the lives and livelihoods of the American people" (National Weather Service, n.d.)

hazard is to occur and the seriousness of the hazard (Griffin et al., 2008). The risk information seeking and processing (RISP) model describes the motivations for information seeking, based on the heuristic or systematic information processing that individuals undergo (Ahn & Noh, 2020). RISP, though stemming from a dual-processing model, relies more on the systematic processing route than the heuristic one. In fact, the RISP model has high explanatory power for outcome variables like information seeking and systematic processing (Yang et al., 2014). RISP has been used in more than a dozen studies, guiding research on environmental, health, and technological risks (Yang et al., 2014).

One of the key components of RISP is risk perception; most RISP studies define risk perception as the combination of perceived probability and perceived severity (Ahn & Noh, 2020; Griffin et al., 2008; Yang et al., 2014). Risk perception is defined as an individual's estimate of the likelihood that a given hazard will cause harm to themself, to others, or to the environment, as well as their perception of the potential severity of that harm (Griffin et al., 2008). In the RISP model, risk perceptions are measured as likelihood multiplied by severity, where individuals score the likelihood of damage in the next year on a scale of 0 to 100, and this number is multiplied by the seriousness of the hazard scored on that same numerical scale (Griffin et al., 2008). The positive relationship between risk perception and negative emotions, such as anxiety and fear, has been thoroughly studied as well through RISP (Ahn & Noh, 2020; Griffin et al., 2008; Yang et al., 2014). However, as systematic processing takes more cognitive effort than does heuristic processing, this model for assessing risk perceptions may not be the most accurate for all situations, especially in cases where judgments are made quickly, like in emergencies.

One heuristic process that has been used to study risk perception is the affect heuristic; it describes how people rely on their initial feelings (positive and negative) associated with a hazard to gauge their risk (Finucane et al., 2000; Siegrist & Árvai, 2020; Slovic et al., 2005). Also known as risk-as-feelings, this approach refers to how individuals make fast, instinctive reactions to a risk based on the "goodness" or "badness" they feel, known as affect (Slovic et al., 2005). Relying on affect is easier than relying on judgment, as it is quicker for someone to assess their initial feelings about a risk than to weigh the pros and cons or pull previous experience from memory. The risk-as-feelings hypothesis postulates that when in a risky scenario, people's decisions are influenced by their anticipated emotions from potential consequences (e.g., am I scared of what happens if this hazards unfolds?), their subjective probability of the risk unfolding (e.g., how likely do I think this threat will actually happen?), and other factors, such as the immediacy of the threat or their underlying mood (Loewenstein et al., 2001). When a decision maker feels negative emotions like worry, fear, dread, or anxiety when making a decision, the resultant choice may diverge from the choice that would have been made under a more rational, cognitive process. In other words, when emotions are factored into a decision, the most logical choice may not be selected.

Emotional reactions influence decisions both the first time someone encounters a risk, as well as if they face a similar risk again (Loewenstein et al., 2001). These emotions can serve as a shortcut for processing risk. Different affective states have different influences on decision-making (Slovic & Peters, 2006). The choice someone made when they were fearful during a past hurricane, for example, would be different than the choice they would make if they were angry; someone who was afraid from past experience may evacuate earlier for an impending hurricane than someone who was mad that they had to evacuate in a previous season.

The last approach to risk perceptions that will be discussed is psychometrics. This is one of the earliest approaches to studying risk and is still practiced. Psychometrics is a blend of psychology and statistics; it refers to the science of measuring psychological attributes, such as risk tolerance (Roszkowski et al., 2005). In psychometrics, risk tolerance is conceptually defined as the perceived risk of an activity compared to its actual risk, which are weighed by the individual against their perceived benefits from engaging in that activity (Fischhoff et al., 1978). What separates psychometric assessments from other approaches to risk perception is that this approach measures acceptability instead of likelihood, which may explain the terminology difference between risk perception and risk tolerance. Regardless of the hazard happening, psychometric assessments look at how acceptable the consequences would be in that future scenario.

The psychometric paradigm was created to make a taxonomy of hazard risk, portraying people's aversions or indifferences towards different hazards (Slovic, 1987). Depicted as a matrix, the psychometric paradigm shows where individuals' ratings of risks fall, ranging on one axis from how dreaded a risk is to how acceptable the risk is on the other axis (Siegrist & Árvai, 2020). These two dimensions (dread and acceptability) mirror heuristic and systematic information processing, as dread is an emotion or affect and acceptability is determined through more complicated processes in memory. However, the psychometric paradigm has focused more on hazard characteristics, whereas approaches like RISP or risk-as-feelings have incorporated more individual characteristics (such as trust in technology or decision-making competence) that can influence risk perceptions (Siegrist & Árvai, 2020). Individual characteristics are often left out of the calculations of risk, due to the complexity of analysis (Siegrist & Árvai, 2020). Yet, demographic characteristics, psychological traits, value orientations, and levels of domain-specific knowledge and understanding have been shown to influence risk perceptions and should be studied more explicitly (Siegrist & Árvai, 2020). The psychometric paradigm is an interesting approach to studying risk; however, it is not relevant to this study because so much of the processing of hurricane risks deals with likelihoods.

Out of the three approaches to think about risk, RISP is the most applicable to this study. Due to project constraints from different interdisciplinary team goals for the wind exceedance graphic, this dissertation is driven by parts of the theoretical RISP model rather than the whole. Individual characteristics and hazard experience, for example, are two of the major inputs of RISP and directly apply to this study. RISP predicts that demographic and individual characteristics influence other variables, such as risk perceptions, affective responses, and information-seeking behaviors (Dunwoody & Griffin, 2015; Griffin et al., 2013; Yang et al., 2014; Yang & Kahlor, 2012). Demographic and socio-cultural characteristics such as gender, race/ethnicity, age, and education have been found to influence risk information seeking and processing (Yang et al., 2014). Extended to this study, many of these demographic categories have also been classified as aspects of vulnerability (Centers for Disease Control and Prevention, 2022; Peacock et al., 2011). Therefore, this dissertation captures individual characteristics in RISP in two methods through the survey responses with demographic information, as well as from experts' metacognitions of the public's vulnerability in interviews. Additionally, it can be argued that individual cognitive abilities, such as a numeracy or spatial cognition (further elaborated on in Section 2.2) also fall under the "individual characteristics" category in the RISP model and thus can influence risk perceptions and decision-making as well. While education level is often used in RISP as an input individual characteristic to explain how people engage with information, numeracy and spatial cognition are more specific measures that can reveal more about peoples' information processing for hazard risk maps. Therefore, this study focuses on individual characteristics to better explain people's abilities to interpret the wind exceedance graphic, as well as their risk perceptions of the hazard.

Previous experience has been found to be a major influence in the RISP model on outcome variables such as information seeking (Yang et al., 2014). Experience with a hazard influences a person's information insufficiency, meaning the gap between what an individual knows about the risk and what an individual thinks they needs to know about the risk, which can impact their future information seeking (Dunwoody & Griffin, 2015). Information insufficiency falls under motivation in RISP to explain what drives people to seek additional information about a risk. General experience in RISP has been defined in studies that focus on health and environmental risk to refer to relevant hazard experience (e.g., Griffin et al., 2013; Yang et al., 2014). Prior hurricane experience was a large focus of this study, as an individual characteristic, because of its influence on decision-making. RISP focuses on information-seeking behaviors, but it does not extend to other behaviors, such as evacuation. This study looks into how to best present hazard information so people continue to process the wind exceedance graphic. Therefore, the input aspects of the RISP model are important to guide this study, though the outcome variables (e.g., accuracy, preference, etc.) are different in order to meet the goals of this study.

2.1.1 Experience and Risk Perception

While studies on RISP have investigated how prior experience influences peoples' affective responses to the risk and ultimately how they seek additional risk information to adjust their information insufficiency, experience has also been measured directly as an influence in risk perception. Prior experience with a natural hazard has a strong influence on people's risk perceptions of the same hazard in the future because of familiarity (Dillon et al., 2014; Houston et al., 2019; Lindell & Perry, 2012). For example, when a storm causes moderate or heavy damage to someone's home, for example, that person is more likely to take a preventative action to protect their property for the next event (Dillon et al., 2014). In their study of the general public, Lindell and Perry (2012) found that when presented with a threat like a hurricane, people assessed their risk by considering their previous experiences, proximity to the storm, and the storm's characteristics (e.g., magnitude). If their level of risk was higher than desired, they looked for the most beneficial mitigative behavior they could do (Lindell & Perry, 2012). While people often decide to take protective actions, they can also decide to not take action to mitigate their risk. Experiencing false alarms and near-misses are one area where even if risk perceptions are high, people may not take action. For example, Houston et al. (2019) found that people who did not experience personal damage during a natural disaster or took protective actions for a storm that failed to manifest had lower risk perceptions for future storms (Houston et al., 2019).

Experiencing a disaster can come from personal experience, as well as from professional experience. Personally experiencing a natural disaster has been shown to affect the risk tolerance of people who work in risk assessment (Bernile et al., 2021). In their study, Bernile et al. (2021) focused on hedge fund managers because this is a group of professional financial risk-takers; hedge fund managers need to make optimal investments for their clients while simultaneously anticipating extreme events that could impact their clients' cashflow. The authors found that hedge fund managers who personally experienced a natural disaster in a given year made significantly less risky investments for their clients in that year, even if their clients were not in a disaster-impacted area (Bernile et al., 2021). The relationship between financial risk taking and experience with natural disasters has been studied in both professional and household (i.e., general public) contexts; results showed that experiencing a disaster or even a near-disaster led to less risky decision-making behaviors both at work and at home (Bourveau & Law, 2021; Gao et al., 2020).

To a limited degree, the relationship between personal experience and risk taking in professionals has been studied in the weather domain. Evidence suggests there is an effect of personal experience on risk perception during severe weather (e.g., Bukvic et al., 2015; Demuth, 2015), but evidence on the influence of professional experience is inconclusive at this point. For example, through in-depth interviews of people from different professional sectors in the Netherlands, Vasileiadou and Botzen (2014) found that individuals who had personally experienced an extreme weather event (e.g., hurricanes, flash flooding, tornadoes) had a significantly higher level of concern about that extreme event than those who did not have this experience. However, people's professional experience with extreme weather events (i.e., those whose jobs required them to work under extreme weather conditions) did not have a statistically significant influence on their levels of concern for those extreme weather events (Vasileiadou & Botzen, 2014). This non-significant result could have been because of the sample size; the authors reflected in their discussion that they felt their non-random sample of 40 individuals was small. The authors found descriptively that professionals whose jobs centered around managing risks and were responsible for others (e.g., crisis managers or public health officials) felt more concerned about extreme weather than those in professions who were out in the elements, making them vulnerable to risks from extreme weather who did not have the responsibility to make decisions (e.g., tour guides or bus drivers) (Vasileiadou & Botzen, 2014).

The influence of professional experience on risk perception should be studied more often. While researchers have looked at how experiencing severe weather affects mental health in the general population and in vulnerable communities for a while, it is only recently that the field has begun to investigate the impacts on professionals in the weather sector whose job responsibilities involve engaging with the hazard at a depth and duration that is longer than would be required if they did not have these careers (Bolton & Ault, 2020; Shukla, 2016). Professional experience could be argued as being even more frequent and recent in the minds of people as they work most days a week, possibly exerting a greater influence on their decision-making (Dittmar & Duchin, 2016). In this study, analyzing professionals like EMs and meteorologists is an interesting extension to these findings on experience, as these groups use severe weather forecast information in their jobs, which is different than the general public and could affect their risk perceptions and decision-making. Testing if professional experience is related to professionals' risk perceptions could provide evidence of an important individual characteristic in risk perception that may impact how hazard graphics are interpreted, potentially extending hazard experience in RISP to professional hazard experience. Thus, the following research questions and hypothesis were posed:

RQ1a: How does experience with hurricanes affect risk perceptions in experts (EMs and meteorologists)?

RQ1b: How does experience with hurricanes affect risk perceptions in the public?

H1: Regardless of group, more hurricane experience will lead to more accurate interpretations of the graphic.

2.2 Universal Design in Hazard Visualizations

Universal design is "the design and composition of an environment so it can be accessed, understood, and used to the greatest extent possible by all people, regardless of their age, size, ability, or disability" (General Services Administration, 2020). Universal design (UD) is essentially good design; products made with universal design in mind are easier for users to employ, are appropriate for a wide range of users with different abilities, and communicate information effectively and equitably. UD started as a movement within architecture with the goal of making physical spaces more accessible for those with and without disabilities (Hitt, 2018). The principles behind UD have been expanded upon and adopted into communication, where professional communicators have begun to consider diverse audience needs in their messaging (Hitt, 2018). UD is a social constructivist approach that has been adopted by many disciplines because its focus is on meeting the needs of all users.

There are legitimate criticisms of UD, citing that the approach is too idealistic and unintentionally discounts the unique needs that many disabilities require (Hitt, 2018). However, messages and products made using UD principles are highly accessible in general. Designing messages to be more accessible is a beneficial practice in information design, despite the challenges of achieving total inclusivity. Especially for communicating difficult concepts like probability information for a changing hazard, messages that are designed to be easier to comprehend accurately can only help audiences' interpretations of the information.

Different audiences may require different depictions of risk when they are processing uncertainty in risk communication; a more technical audience would be more likely to understand probability theory, for example, than would a more general audience (Bostrom et al., 2008). While scientists can design different visualizations for each group, often one visualization is disseminated widely. Data visualizations, therefore, should be designed so that expert and nonexpert audiences can effectively interpret the risk from a hazard (Bostrom et al., 2008). For the new wind exceedance graphic, universal design can be applied to three parts: probability language, color schemes, and localization through cartographic landmarks. These topics are discussed below in more detail.

2.2.1 Probability Language

Probabilities are one way to quantify uncertainty in data (Johnson et al., 2019). Mathematically, probability describes the likelihood an event will occur between 0 and 1, where 0 indicates something will not occur and 1 marks absolute certainty. In meteorology, probability is commonly used to depict the percent chance of weather hazards occurring for a time period. Despite that probabilities are commonly used, forecasters have voiced concerns about how understandable probability expressions are for different user groups (Ripberger et al., 2022). Interpreting probability information can be difficult for both experts and the public; accurate interpretations can be further affected by the format in which the probability information is communicated (Ripberger et al., 2022).

Probability is a common, anticipated feature of meteorological forecasts. In storm forecasts, for example, people expect weather information to be presented using probabilities, and they then process their risk from the storms in probabilistic terms (Joslyn & Savelli, 2010; Morss et al., 2008). Risk perceptions and probability language have been studied in tandem, as the form in which a risk is presented can have a major effect on how it is perceived. For example, researchers have studied how different presentations of probability affect people's understanding of uncertainty information. In a study of how EMs interpret tsunami evacuation graphics, Lindell et al. (2021) found that probabilistic language, such as 0.10 or 10%, better conveyed forecast uncertainty than did ratios, e.g., "1 in 10." EMs misinterpreted "1 in 10" more often than percentages because it was unclear what the "10" represented, even though the two expressions are numerically equivalent (Lindell et al., 2021). Another study found EMs rated risks for damage from a storm to be more unacceptable when it was reported in the “1 in 10” format than when the risks were reported as a 10% probability (Wernstedt et al., 2019). Thus, understanding more about how people understand percentages and ratios could provide insights into the best way to present this information for various audiences.

Probability can act as a heuristic cue in decision-making (Tversky & Kahneman, 1993). For example, if given the percent probability of a hazard (such as a 40% chance of rain), people can make a decision based on their mental shortcut of what the number has meant in the past rather than what the number actually means. While 40% is a lower probability, if it typically rains when this percentage is forecast, then people may assume that it always rains when the forecast is 40%. Taking a mental shortcut for interpreting probability can work for both numbers and verbal descriptions of probability. A slight chance of rain means that there is a 15% to 24% chance of rain in a given time (Ripberger et al., 2022). However, someone who interprets a slight chance of rain to mean a 10-20% chance would likely interpret a “slight” chance of a severe storm (e.g., a hurricane or tropical depression) to be higher (e.g., 20-30% chance) even though the two events have a similar probability of occurring (Ripberger et al., 2022). Even though “slight” is the same word, someone’s perception of a more severe storm can alter how they think about “slight” as a modifying phrase for the probability. This difference is important for forecasters to consider when they use verbal probability statements, as different numerical interpretations of the same words and phrases vary based on the type of storm forecasted could impact what protective actions people take when under a watch or warning.

One area that has been getting attention within the weather sector is numeracy, meaning a person’s ability to understand or work with numbers, like probability (Ripberger et al., 2022). On average, the public and EMs are less numerate than meteorologists, which impacts the public’s and EMs' comprehension of forecast graphics that contain probability information (Reinhart et al., 2021). Designing graphics so that people with lower numeracy understand them has become an increasing practice in the weather sector in order to have the greatest likelihood of user groups accurately comprehending the forecast (Ripberger et al., 2022). For instance, studies indicate that graphics with probability information should use both text and numbers to assist those with lower numeracy (Lenhardt et al., 2020; Reinhart et al., 2021; Rosen et al., 2021). For example, instead of using graphics titles that state "10% Probability of Wind Speeds Exceedance" or "Low Probability of Wind Speed Exceedance," use "Low Probability (10%) of Wind Speed Exceedance."

Numeracy has been argued to be an important skill that all adults should have, as the ability to understand and accurately interpret numbers is necessary for everyday financial and health decision-making (Peters et al., 2019). Numeracy is often measured as a mathematical ability to process and understand basic probability and numerical concepts (Peters et al., 2006). However, numeracy is more nuanced than just an objective ability to work with numbers; there is also numeric self-efficacy, which describes one's confidence in their abilities to use numbers (Peters et al., 2019). Having higher numeric self-efficacy was found to relate to higher engagement with numerical information and higher completion rates for mathematical tasks (Peters et al., 2019). As ability and efficacy exist in a positive feedback loop, it makes sense that confidence is an essential component of numeracy; numeracy depends on a person's ability to both accurately and confidently use numbers (Peters et al., 2019; Scott, 1999). Therefore, this study captures both numeracy and confidence to examine the relationship in this definition.

Many of the studies that look at numeracy *and* confidence come from the education discipline (e.g., Campbell et al., 2020; Ferme, 2018). In the forecasting and disaster literature, numeracy has been analyzed in a fair number of studies of EMs and meteorologists, though not in conjunction with confidence (Adams et al., 2017; Ripberger et al., 2022). Thus, this study aims to extend the relationship between numeracy and confidence by investigating it in the context of hazard graphic interpretation for expert user groups and the public. This study predicts relationships around numeracy, confidence, and accuracy of interpretation. The following three hypotheses test numeracy where it includes both confidence and accuracy:

H2: Experts (EMs and meteorologists) will be more accurate in interpreting probability information in the wind exceedance graphic than will a public sample.

H3: Experts (EMs and meteorologists) will feel more confident interpreting the wind exceedance graphic than will a public sample.

H4: Numeracy and confidence will have a positive relationship.

2.2.2 Color Choice and Texture

Weather forecasts and hazard maps are notorious examples of data sets that are highly relevant to much of society, yet are repeat offenders of using rainbow color schemes (Crameri et al., 2020). Rainbow color schemes are difficult for those with protanopia (red-green colorblindness) to see accurately, especially if the color saturations in the spectral scale are the same, regardless of hue (Light & Bartlein, 2004). Mixing color temperatures, meaning how warm or cool a color is, can be similarly difficult for those with protanopia, as red is on the warm color spectrum and green is on the cool color spectrum (Crameri et al., 2020; Light & Bartlein, 2004). Especially for hurricane map graphics, rainbow schemes render the hazard information difficult for colorblind people to read accurately, which can put these individuals at a higher risk (Crameri et al., 2020). While globally only about 0.5% of women and 8% of men have a color-vision deficiency, using colorblind-friendly schemes will only increase the accessibility of a graphic, which is an important tenet of UD (Crameri et al., 2020; General Services Administration, 2020). Color schemes that accommodate protanopia accommodate most other forms of color vision deficiency (Light & Bartlein, 2004).

Color schemes can have data distortion effects when used ineffectively (Crameri et al., 2020). Using colors that have unequal saturation variations, for example, can signal to the user that there are different weights to the data being presented, even if that is not true (Crameri et al., 2020). Additionally, there are psychological color associations that should be considered when colors are connected to data. Research indicates that red is associated with the presence of a hazard and blue with the absence of a hazard (unless for flooding, where blue is associated with the presence of water) (Thompson et al., 2015). Additionally, red often communicates high risk or danger (Bostrom et al., 2008; Cheong et al., 2020).

The only time that color choice has been directly analyzed in the development of a forecast graphic was for storm surge. Storm surge forecast graphics were some of the first to have interdisciplinary, rigorous testing of prototype maps, and this work has been an inspiration for this study. Morrow et al. (2015) wanted to find out the preferences and perceptions of storm surge communication materials from experts (forecasters, broadcast meteorologists, and EMs) and the public. To answer their questions, Morrow et al. (2015) initially assessed interpretations of and preferences for inundation maps using one-on-one interviews, focus groups, and webinars. They used the results from their qualitative research to determine how color, legend text, and landmark use should be altered for the quantitative testing planned in a series of surveys. The research team send the same survey out over three years, starting in 2011, to capture the public's perspective. After the surveys were collected, Morrow et al. (2015) quantitatively analyzed risk perceptions for storm surge in tropical and extratropical cyclones.

The analysis from all three years found that there was support for having separate storm surge warnings, with unique text and colors in graphics.

The current storm surge graphic product uses a blue-yellow-orange-red color scheme, where each hue has the same saturation value. While this approach is better than a rainbow color scheme, it still does not follow UD. Thus, I am testing two red-based color schemes for the wind exceedance product, as shown in Figure 3.

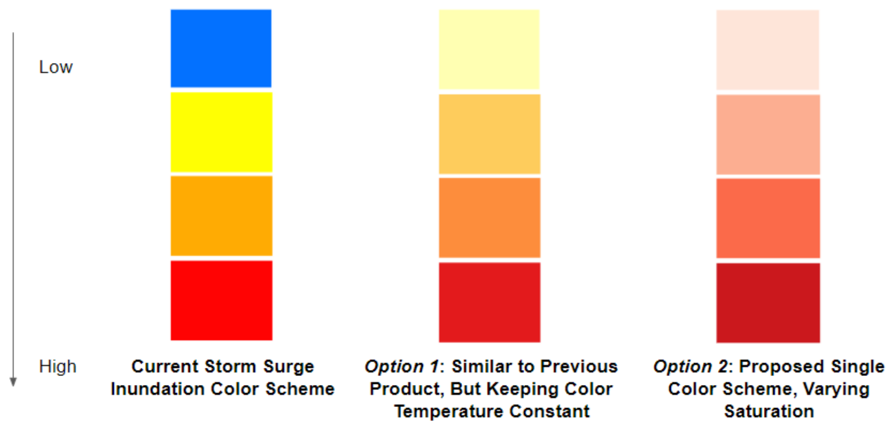


Figure 3. Proposed color schemes for the wind exceedance product

Of the proposed color schemes shown in Figure 3, option 1 is the most similar to the current storm surge graphic; it differs from that graphic in that it uses yellow as lowest category instead of blue. This substitution serves two purposes: a) it lessens the chance that the wind exceedance graphic will be confused with the storm surge graphic, and b) it results in a more colorblind-friendly graphic since the color temperature would be constant (i.e., only warm colors). Option 2 only uses red, showing a single-color option that follows the recommendations from the emergency graphics literature and has a more obvious variation in saturation between categories. Testing both schemes is important since many graphics in the weather sector use the rainbow color scheme or multiple colors to delineate categorical information, and the use of one color with varied saturations could have a mixed reception from users.

Color hue and texture are among the most common principles of cartographic symbolization (Cheong et al., 2020; Lindell et al., 2021). Thus, this study tested not only two new color schemes with varied hues that use the color red (Figure 3), but it tested the use of varied textures for the exceedance graphic. Texture describes a representation on a map that uses a pattern instead of color to show the difference between categories. An example of using texture is shown in Figure 4.

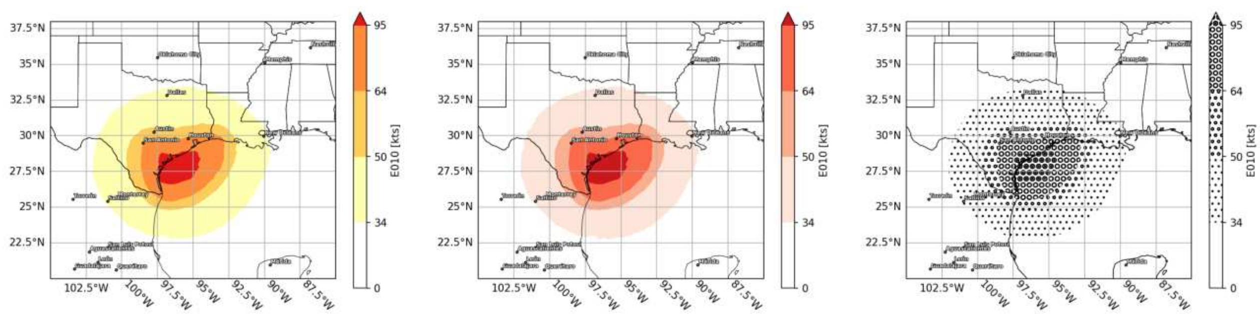


Figure 4. Examples of color and texture schemes for the wind exceedance graphic prototypes

Texture has been used in forecast track visualizations, known as the cone of uncertainty. The cone of uncertainty displays the most likely path of a hurricane, determined based on a five-year sample of past hurricane tracks while accounting for two-thirds of official forecast errors (National Hurricane Center & Central Pacific Hurricane Center, 2022). The cone is solid in color for the 3-day projection of a storm track and is shaded to show the 4- to 5-day forecasted track (shown in Figure 5). This visualization is one of the most recognized hurricane graphics by the general public, though it is consistently misinterpreted (Broad et al., 2007; Demuth et al., 2012; Witt & Clegg, 2021).

The shaded pattern in the cone for days 4 and 5 forecasts has been interpreted to mean that the forecasted path is more uncertain, whereas the shading is meant to convey the difference in model certainty on these days (Eosco, 2008; Ruginski et al., 2016). Texture has not been used in hurricane graphics to communicate category (e.g., wind speed thresholds) Thus, this study uses texture as an option to convey categorical differences in information, as this scheme could improve the readability of the graphics if it is not misinterpreted as uncertainty.

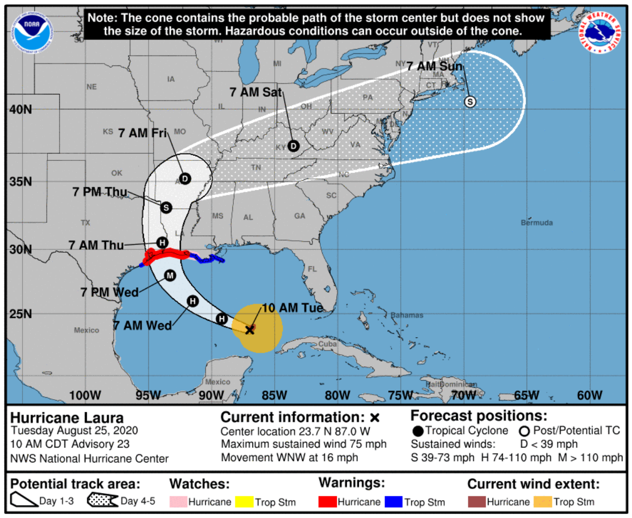


Figure 5. Cone of Uncertainty sample graphic (National Hurricane Center & Central Pacific Hurricane Center, 2022)

Texture has been recommended as one of the best options for representing uncertainty in graphics because it has a more universal visual separability between categories than color schemes (Bostrom et al., 2008; Maceachren et al., 2005). That is, texture is a colorblind-friendly option since the variation between categories is based on the visual density of the pattern, rather than a change in color. Additionally, as there would not be any psychological color associations with the graphic, texture could be an accessible design application for hazard maps.

Studies of texture representations have produced mixed results. For example, Cheong et al. (2020) found that texture representations in an emergency graphic resulted in EMs making decisions that were less risky, yet slower (up to 7.8 seconds slower on average) when compared to graphics using color schemes with varied hues. The texture option was also the EMs' less preferred scheme.

Gathering more information about user preferences is valuable, in addition to learning how color and texture are perceived in risk maps. The findings could inform the development of future forecast products. To that end, I propose the following questions and predictions:

RQ2a: What are experts' preferences for the color scheme and texture in the wind exceedance product and why?

RQ2b: What are the public's preferences for the color scheme and texture in the wind exceedance product?

H5: Regardless of group, users will perceive the reds-only color scheme as riskier than the yellow-to-red color scheme.

H6: Regardless of group, there will be a relationship between how risky the users perceive the colors schemes and which color scheme they prefer.

2.2.3 Localization

Localization in maps describes the adding of local contextual features, like identifiable places or landmarks (such as roads, municipal buildings, names of neighborhoods), to help an individual visualize the spatial extent of a hazard (Henstra et al., 2019). Using localization in product design could promote more effective communication of the hazard information, which is a goal of UD.

Localization has been discussed in the context of Hurricane Local Statements (HLSs), which are weather statements made for local, affected communities during incoming hurricanes and tropical storms. Some broadcast meteorologists have reported that they would like even more information about local impacts than what is currently in a HLS (Reinhart et al., 2021). EMs have also said they would like more local detail on storm surge maps for planning and operations purposes (Reinhart et al., 2021). Thus, designing graphics using design principles that encourage localization should benefit the expert users (e.g., EMs and meteorologists) who will use the wind exceedance product most often.

One cartographic method to promote localization in map design is to add major cities. People are drawn to their own location when viewing maps because of the personal relevance, so including large, nearby reference cities should make it easier for them to find their location (Lindner et al., 2018). When people read a hazard map, their attention will first be drawn to their personal location on the map, and then they will focus on the information about hazards being communicated in the map (Cheong et al., 2020). Using city landmarks on the new wind exceedance graphic could assist users in focusing on their location and interpreting the hazard information using a more personal, localized context. Examples of including landmark cities in the wind exceedance prototypes are shown in Figure 4, as well as larger below in Figure 6.

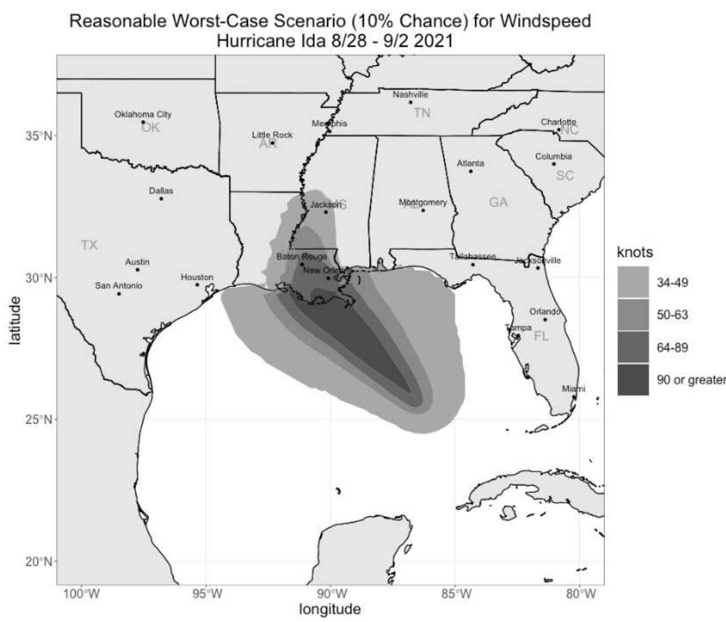


Figure 6. Close-up of landmarks on wind exceedance graphic

Maps that display hazards at even more local levels than cities, like a block or neighborhood, could be more engaging for map users since the risk is shown alongside items that make the area more familiar to them, like roads (Monmonier, 2008; Retchless, 2014). Evacuation maps are one example of hazard maps with imposed road networks to make the maps more useful (Monmonier, 2008). In an assessment of online sea-level rise maps, Retchless (2014) found that when the threat was clearly emphasized as a local threat, the impacts became more personally relevant to the user. Sea-level rise is tricky to personalize, as it refers to climate change, which is a longer-term threat that may not be witnessed in a community during one's lifetime. In digital maps of sea-level rise, however, depictions of flooding over familiar places and roads made climate change more obvious, signaling to the audience that the risks of climate change should be assessed instead of dismissed (Retchless, 2014).

NOAA has added roads on its maps to communicate about storm surge, but the effect of this localization has not been assessed. Thus, studying the addition of roads or city landmarks on the wind exceedance graphic could reveal the types of localization preferred by different user groups in forecast graphics. The following research question and hypothesis guide the analysis on incorporating overlays onto the map graphic:

RQ3: Are cities or roads more preferred as landmarks for localizing the wind exceedance graphic by experts and the public?

H7: Regardless of group, users will have higher perceived riskiness of the wind exceedance graphic with landmarks than for the graphic without landmarks.

2.2.3.1. Spatial Cognition

Spatial cognition can be defined as the ability to orient oneself within a map and comprehend its purpose (Tomaszewski et al., 2020). Similar terms like spatial awareness or spatial thinking ability are used in the cartography, emergency management, and forecasting disciplines to describe how people familiarize themselves in space and draw conclusions based on where items are located in a map (Alexander, 2004; Hegarty, 2010; Houston et al., 2019;

Kane, 2014). Thus, this review discusses studies that analyze spatial cognition, using the terms the authors chose for their work. In this study, I use "spatial cognition" to mean a person's cognitive ability to understand the location of items in space when shown in maps.

Using maps, Houston et al. (2019) studied the influence of spatial awareness on people's risk perceptions of flash flooding. They found a difference between participants' flood risk perceptions, such as awareness of where flooding was likely within a community, and their nonspatial dimensions of flood risk perceptions, such as general dread or non-geographically specific concerns about flooding impacts. Length of residency in a flood-prone area was positively associated with spatial flood risk perceptions, though other demographics (e.g., gender, income, age, and educational attainment) did not have a statistically significant association (Houston et al., 2019). Spatial abilities have the potential to be a major influence on how users interpret graphic information and should be considered when discussing the ease of using a map.

The study of spatial thinking ability has extended into the meteorological domain because spatial displays are used by expert forecasters and general audiences (Hegarty, 2010). In a study of weather forecasters in the U.S. Navy (n = 21), Smallman and Hegarty (2007) asked participants to generate forecasts and use complex weather maps for a modeled carrier group. In this study, the researchers measured participants' spatial thinking ability using psychometric tests, such as a mental rotation test and paper folding test. They found that there were differences in accuracy of interpreting the forecast maps, though not statistically significant. When compared to a public sample in a future study, the authors found expert users (Navy forecasters) were more accurate than novice (general public) groups, especially for maps with features like state lines or mountain ranges as landmarks (Hegarty, 2010; Smallman & Hegarty, 2007). However, the Navy forecasters' spatial ability was not statistically different from the general public's sample (Hegarty, 2010). This finding was surprising as one would expect that career forecasters who use maps daily to have a greater spatial ability than a non-specialized person. While this result could be due to a lack of statistical power, the tests used to measure spatial thinking ability could also be a cause. Smallman and Hegarty used computational psychometric tests that measure general spatial ability, whereas they wanted to capture map reading ability, which the tests used do not measure. Spatial cognition, therefore, should be measured using a spatial test or scale that is more related to using maps, such as those from the navigation or sense of direction scholarship (Hegarty et al., 2002).

Relationships between levels of spatial cognition and numeracy have been found as well, though these concepts are not often measured together (Kane, 2014). Decision-making often involves numerical assessment as well as a consideration of the local context, so including spatial cognition with numeracy could elaborate on different individual characteristics that influence how choices are made for a given situation in a community (Dowling, 1997). In their study of recycling and refuse operators, Kane (2014) found that the operators considered spatial factors when making calculations as part of their job, e.g., measuring available bin space for collection when planning the day's pick-up schedule. Additionally, knowledge of roads assisted operators in calculating the number of pick-ups they needed to make to earn a profit over their operating costs, as they had to consider things like traffic patterns, time of day, and which neighborhoods typically had the heaviest bins (Kane, 2014). Connecting numeracy with spatial cognition should be extended more into other professional domains, like emergency management and forecasting.

In emergency management decision-making, an EM's geographical knowledge is made of environmental knowledge (e.g., where floodplains or spots of coastal erosion are located) and knowledge of infrastructure (e.g., roads) (Alexander, 2004). However, this geographical knowledge is not innate, but learned. EMs develop their ability to interpret maps over the course of their education and careers. One of the major conclusions drawn in a study of emergency management students by Alexander (2004) was that developing spatial cognition is a critical skill for operational career professionals who manage disasters and crises, along with other skills such as communication and decision-making. Geographic information systems (GIS) training often facilitates the development of spatial cognition, and is required in the curricula of emergency management and atmospheric sciences programs (Tomaszewski et al., 2020). Just as the literature shows that numeracy influences confidence, spatial cognition may predict how people accurately interpret forecast information from maps. Predicted relationships that arise include:

H8: Numeracy and spatial cognition will be higher in the expert group (EMs and meteorologists) than a public sample.

H9: Numeracy and spatial cognition will predict the accuracy of interpretation of the graphic.

2.3 Vulnerability and Decision-Making

Vulnerability describes the social and environmental factors that can make a population more at-risk to suffer from a hazard (Oliver-Smith & Hoffman, 2002). There are ecological, political-economical, and socio-cultural dimensions to vulnerability (Oliver-Smith & Hoffman, 2002; Wisner et al., 2004). Vulnerability can be more concretely defined as a "characteristic of a person or group and their situation that influences their capacity to anticipate, cope with, resist, and recover from the impact of a natural hazard" (Wisner et al., 2004, p.11). Class, occupation, caste, race/ethnicity, gender, disability and health status, age, social network, and citizenship status are all key variables to consider in assessing a group's vulnerability. Cumulative vulnerabilities need to be considered in emergency management, such as the combination of racial, economic, and environmental factors that make a group more vulnerable.

In hazard scholarship, some definitions of risk (the probability of a hazard) have included more expansive definitions for hazard, where "hazard" includes the probability of the event happening as well as the event's potential magnitude of impact on society, the environment, and the sociopolitical contexts therein (Whyte & Burton, 1980). During natural disasters, the susceptibilities of socio-politically/socio-economically vulnerable communities interact with the vulnerabilities in the physical geography, impacting different populations at different levels (Peacock et al., 2011). For example, during hazardous weather, lower-income communities are at greater risk because they are often situated in low-lying areas and in lower-quality homes (Peacock et al., 2011).

Previous studies have looked at coastal communities in hurricanes through the lens of vulnerability (e.g., Bathi & Das, 2016; May, 2019; Niles & Contreras, 2019; Santos-Hernández, Méndez-Heavilin, & Álvarez-Rosario, 2020). In their vulnerability assessment of communities in Mississippi, Bathi and Das (2016) presented a quantitative framework to measure both socioeconomic and flood inundation (environmental) vulnerability. The socioeconomic vulnerability indicators adopted in their study were: total population, non-white population, number of female-lead households, population under 18 years old, population 65 years and older or disabled, number of households without a vehicle, number of families in mobile homes, number of people in group quarters, number of people below poverty in the past 12 months, population 18 years and over with no diploma, and number of households with no one age 14 and over who speaks English well (Bathi & Das, 2016). A community was considered to be environmentally vulnerable if the census showed that it was located in a 100-year flood plain (Bathi & Das, 2016).

Social vulnerability describes “the characteristics and diversity of populations in terms of broader social, cultural, and economic factors that shape abilities to anticipate future events, respond to warnings, and to cope with and recover from disaster impacts” (Peacock et al., 2011, p.4). Coastal communities that are exposed to hurricane hazards are particularly vulnerable, as these areas may not have the same resilience capacity season after season (Peacock et al., 2011). Scholars in emergency management have done an increasing amount of work around learning how to adapt different disaster prevention and recovery strategies to the realities faced by vulnerable communities (Rubin & Cutter, 2020). Asking EMs and meteorologists about the decisions they make for these communities, such as when they would mobilize resources or when they would open evacuation centers, could reveal how these expert user groups take into account the vulnerabilities in their community to better protect people.

Metacognition describes our feelings and beliefs about our thoughts and actions in any given scenario (Huang & Yang, 2020; Kimrey, 2017). Community leaders and decision-makers often use their metacognitions of their decisions from past experiences to think through complex situations that can impact themselves and their constituencies. During crises, like oil spills or major hurricanes, EMs rely on mental shortcuts from past crises to make quick decisions in a new crisis to help keep their vulnerable communities safe (Kimrey, 2017). For example, in a current storm, EMs may decide to close certain roads they have closed in past hurricanes because they remembered how effectively those decisions played out in the past and how vulnerable different populations were along that route. While each hurricane follows a different track, the mental shortcut of knowing what roads to close or communities to evacuate because they have more vulnerabilities can make EMs more effective decision-makers. EMs can use their reference knowledge for how they thought through decisions they made during a hurricane in the past to inform their current actions. If EMs are aware of these shortcuts, they can work on building their metacognitive abilities towards decision-making in future crises. Likewise, once these metacognitions are known, other community partners like meteorologists can create forecast graphics that are more useful to EMs when they are making fast, impactful decisions. For example, if we know evacuation is thought about by EMs as road networks between vulnerable, coastal areas to safer inland routes, meteorologists can emphasize which highways are the most critical to highlight during briefings.

Many theories fold metacognition into crisis communication, including that of meta-leadership. Meta-leadership combines the theories of metacognition and organizational leadership to describe how individuals think about their leadership during emergencies, where the role of being a leader often requires coordination and communication amongst diverse agencies and sectors (Marcus et al., 2006). One of the reasons meta-leadership theory was developed was in response to the coordination failures between agencies in the aftermath of Hurricane Katrina; these failures amplified the destructive impacts of the wind and levee flooding on the residents of New Orleans (Marcus et al., 2006). Emergency managers are meta-leaders, as they have to think through risk management from multiple perspectives in order to make decisions. Meta-leaders also have to manage uncertainty by making choices when faced with a dynamic hazard or threat (Kimrey, 2017).

Uncertainty has been conceptualized as a metacognition for risk perceptions (Huang & Yang, 2020). How confident individuals feel about a risk judgment they have made is a metacognition of their feelings of uncertainty (Huang & Yang, 2020). Meta-leaders need to consider many perspectives when enacting emergency response plans. As these leaders need to make decisions while they parse through uncertainty information, understanding how meta-leaders take into account different populations' needs could reveal how they think through uncertainty when assessing their community's overall risk. In this study, my goal was to see how vulnerability factors into the metacognitive decision-making that comes from using the data visualizations in forecast products. Vulnerable populations do not typically have the same access to information and resources to cope with a hazard warning (Peacock et al., 2011) that more advantaged populations have. Thus, learning how EMs and meteorologists consider these populations when they are making decisions in a hurricane event could provide context for how these meta-leaders use forecast graphics. A formal question that arises:

RQ4: What metacognitions do experts (EMs and meteorologists) have of vulnerability for hurricane-prone areas?

CHAPTER 3. METHODS

The following methods section covers the study's three phases: (a) graphic development, (b) testing with expert user groups, and (c) testing with a public sample. This study used a mixed-methods approach by combining qualitative and quantitative analyses to answer the research questions and hypotheses. Mixed methods approaches are useful to compare within and across types of questions, as well as for elaborating on the experiences and perceptions of individuals and groups more broadly (Frels & Onwuegbuzie, 2013). The methods for each phase are detailed in the sections below.

3.1 Graphic Development

During the development phase, I used a combination of best design practices from a literature review, feedback from monthly inter-agency meetings, and informal interviews with key stakeholders from the NHC and the Federal Emergency Management Agency (FEMA) to develop the prototypes of the wind exceedance graphic.

I created the earliest versions of the prototype graphic (see Figure 4) based on the best practices identified from the literature in emergency management, meteorology, universal design, and mass communication disciplines. My selected design choices were then put into prototypes[3] that were presented at the 2021 National Oceanic and Atmospheric Administration (NOAA) Hurricane Conference to obtain initial feedback from the operational meteorology community. The conference took place November 30 to December 3, 2021; I attended the

[3] Because the NWS typically uses one graphic to convey a data concept (such as wind speed) to multiple audiences, I tested multiple prototypes for a single wind exceedance graphic.

conference via remote access over Google Meet. At this meeting, Andrea Schumacher (Societal Impacts Lead in CIRA) presented the initial prototypes during a session focused on updates to the forecasting models, and I took notes on the early design impressions, including positive remarks from forecasters and program managers regarding the theoretical backing for the suggested design choices.

At the conference, the CIRA/CSU team made connections with NOAA's Tropical Roadmap Team. The Tropical Roadmap Team consists of members from NOAA's NWS forecasting offices and national centers. The Tropical Roadmap Team researches potential products, services, and policy changes by synthesizing social science findings and recommendations that pertain tropical weather events, and putting their findings into empirical and white papers to help forecasters best use NOAA and NWS products with their intended audiences (Reinhart et al., 2021). The CIRA/CSU team was invited to attend bimonthly Tropical Roadmap Team meetings, which we attended for one year. The aim of these meetings was to provide support for carrying out the project plans to accomplish both the Tropical Roadmap Team's and CIRA's goals. In these meetings, I recorded feedback regarding the wind exceedance prototypes and got access to previous white papers. Graphic design recommendations in these white papers mirrored those from the literature review in this dissertation. These meetings helped to solidify the design of the initial prototypes, as well as set up connections with other stakeholders from the NHC.

During this time, the CIRA/CSU team organized a meeting with NHC forecasters and the branch chief of the Hurricane Specialist Unit to gain more feedback on the wind exceedance prototypes. This meeting was on March 3, 2022. At this meeting, I showed the NHC forecasters the prototypes and took notes on their impressions of the design, as well as the proposed methodology for the testing phase of this project.

The NHC forecasters provided mixed advice on the color schemes for the prototypes. NHC forecasters did not favor storm surge design practices (see Figure 3) when they were used in other graphics, like the time-of-arrival graphic for hurricane winds (Berg, personal communication). I showed the three unique color scheme options for wind exceedance, a yellow-to-red scheme, a reds-only scheme, and a texture option. There was equal support for the yellow-to-red color scheme and the reds-only color scheme. There was an early pique in interest around the textured option, though this was not discussed as much as the color schemes.

After the meeting, the NHC forecasters shared past hurricane briefings prepared by EMs at the NHC with me to provide context for how wind speed probabilities were discussed in the past season. Hurricane briefings are multimedia presentations the NWS forecasters and EMs give to their stakeholders that show graphics, radar imaging, and text about how the storm is anticipated to behave and the potential impacts that may arise. Briefings are given in-person and virtually, as some stakeholders listen to the session via online live streaming or via phone conference calling. From my review of these recorded briefings, I learned that during the 2021 hurricane season, wind exceedance was discussed with the time-of-arrival graphics, but it was only discussed once. I also learned that wind speed probabilities were put into tropical storm briefings more often before watches and warnings had been assigned to the tropical storms than in briefings issues afterwards. The information gathered during the NHC meeting and from watching the past hurricane briefings helped me understand how wind data has been used in an active hurricane center and solidified for me that I wanted to interview EMs and meteorologists to gain an even deeper understanding of how wind probabilities are used by these groups.

The branch chief from the Hurricane Specialist Unit suggested that having more virtual options for interviews could be easier for the desired population of EMs and meteorologists (Hogsett, personal communication). However, some in-person approaches, such as attending briefings in person during the active hurricane season, would provide context on what exceedances were the most useful to EMs and allow for networking between the CIRA/CSU team, current forecasters, and EMs. Thus, it was determined that Andrea Schumacher, a research associate at CIRA who is a lead investigator on the JTTI grant, and I would plan go to the NHC in the 2022 Atlantic hurricane season to observe hurricane briefings where connections for future interviews could also be made. In addition, the Hurricane Specialist Unit group members said they were willing to provide feedback on the graphic prior to the 2022 hurricane season if we wanted more advice.

To gain more context for why the wind exceedance graphic was wanted by the field, I conducted an informal interview prior to hurricane season (July 2022) with Matthew Green, the FEMA Response representative on the Hurricane Liaison Team of the NHC, and Tim Oram, branch chief at NOAA's Operational Services Division. Both Matthew Green and Tim Oram have been involved in producing the wind speed probability models for previous graphic products, and CIRA team members, such as Mark DeMaria (the technology and science branch chief at the NHC and a CIRA fellow), have said these two people were instrumental in pushing for the creation of the wind exceedance graphic. Since wind exceedance has not been displayed in its own graphic yet, knowing why it was recommended that these thresholds be displayed in their own probability graphic suggests how EMs and forecasters are expected to use the product. Additionally, this interview provided insight into other design features that should be included in the next version of the prototype.

3.1.1 Informal Interview with NHC Liaisons

On July 27, 2022, I conducted an informal meeting with Matt Green (FEMA/NHC), Tim Oram (NOAA), Jessica Schauer (NWS), Michael Spagnolo (FEMA/NHC), Eric Blake (NOAA), Andrea Schumacher (CIRA), and Mark DeMaria (CIRA/NOAA) to gain perspective on why wind exceedance as a specific graphic is needed in the hurricane graphic suite of products. In this hour-long discussion, we discussed that having a wind exceedance graphic could be helpful for training expert users. EMs want to know how strong storms could be at landfall, which is tricky to estimate with existing wind models because they have inconsistencies in predicting inland windspeed reductions. At the annual NHC conference the previous year, DeMaria and Oram spoke about how EMs and forecasters were asking for this kind of graphic product, especially for the 90th percentile and 10th percentile (reasonable worst case) with time-of-arrival information.

When asked how they anticipated EMs would use the new wind exceedance graphic, Michael Spagnolo noted that exceedance information is most useful in rapid intensification[4] scenarios. Spagnolo also stated that most EMs would make suggestions to their community partners to enact plans that go into effect for one category higher than forecast. For example, wind exceedance information can be used for deciding when equipment mounts need to be taken down due to the wind and for looking at structure issues and faults.

In wrapping up this meeting, the group expressed concern over how exceedances would be interpreted by non-emergency management audiences, such as the public. While this conversation focused mainly on EM and NWS users for the graphic, Oram was concerned about how the public would use the data and how to design the graphic for them. One suggestion was

[4] Rapid intensification describes a cyclone that speeds up greatly in a short period. The formal definition from the NHC states, "an increase in the maximum sustained winds of a tropical cyclone of at least 30 knots in a 24 hour period" (National Hurricane Center, n.d.).

to use units such as miles-per-hour instead of knots in the graphic, since most people do not use knots in their everyday lives. Overall, this informational interview brought up topics that this study aimed to address, which was encouraging for the study design.

3.2 Testing of the Prototype with Expert Users

In the testing phase, the planned methodologies included a participant observation of live hurricane briefings during the 2022 hurricane season and semi-structured interviews with EMs and meteorologists about the wind exceedance graphic prototypes using simulated storm cases. Hurricane briefings are twice-daily presentations from on-site EMs and lead forecasters about what the current storm conditions are, where the storm track is expected to move, and any anticipated impacts that are forecast to occur. By observing and taking field notes during the hurricane briefings, I would have gained context for how wind exceedances and other wind speed probabilities (e.g., likelihood that a point on a map will be struck by hurricane-force winds, 64 knots/74 mph, during the next 5 days) are discussed during an active threat, as well as whether other community factors (e.g., types of housing along a storm track) are brought into EMs' live decision-making. Additionally, the in-person observation would have provided an opportunity to meet EMs and meteorologists who might participate in the semi-structured interviews.

Participant observation describes the process where a researcher enters a group of people with a shared identity to gain a deeper understanding of the group's actors, interactions, and norms (Blevins, 2017). While the researcher may participate in the group they are joining, observation and notetaking are the key components of this methodology. This method has been increasingly used in organizations or workplaces, as there can be unique cultures that arise within a professional community that warrant study (Blevins, 2017).

To negotiate access to a live hurricane briefing, the CIRA/CSU group reached out to the Hurricane Specialist Unit for permission to observe the materials used and topics discussed by EMs and meteorologists for an unfolding threat at the NHC. Initial permissions were granted from Michael Brennan (branch chief of the Hurricane Specialist Unit), Daniel Brown (senior hurricane specialist and warning coordination meteorologist), and Robbie Berg (senior hurricane specialist) to set up an observation. The observation was to take place ideally during a storm that was *not* forecast to impact Miami, Florida, for the safety and availability of the researchers.

As explained in Section 4.1, observations were planned to consist of attending the twice-daily hurricane briefings for up to a week in August, September, or October of 2022. I planned to take field notes that reflected the atmosphere of the briefing room, the roles and responsibilities of the actors within the room, and the events and interactions that took place between these actors. Participants in briefings at the NHC are accustomed to researchers observing them, so there was minimal concern around this study's impact on the proceedings within the room. However, I planned to use reflexivity to reflect upon my experiences and knowledge of the forecasting domain as I reported my observations. My notes captured in the briefing would have undergone perception checking, which describes the process of checking with group members about whether the perceptions of what is happening match between parties (Blevins, 2017). I planned to send my notes to a couple of EMs who attended to see if they agreed with my observations, as well as to Andrea Schumacher who would have attended the briefings as a researcher. Data from the participant observations would have informed the semi-structured interviews; for example, information learned about how wind speed probabilities are used in briefings could have informed how I asked about risk perceptions for the hazard as well as potential decision-making stemming from the information. Additionally, seeing how graphic products are displayed in briefings could have helped to make our simulation with the prototypes more realistic in the interviews later.

After I gained connections and context for how wind speeds are discussed by EMs and meteorologists, I planned to conduct semi-structured interviews within a testbed. Though these interviews did not happen within a testbed, in a testbed, new products and messages are brought to operational forecasters and other active partners for testing and feedback. Testbeds also allow researchers to observe these professionals as they make decisions during active storm days and simulations and to assess what products are the most relevant in these environments. Testbeds are similar to usability testing; they are conducted either as focus groups or as individual interviews. Researchers and forecasters are brought to a NOAA research site to work side-by-side, evaluating emerging research concepts and tools in simulated operational settings (Tarp, n.d.). The Hazardous Weather Testbed (HWT) is one of the more established testbeds, where as many as 60 researchers and forecasters are brought to the NOAA Severe Storms Laboratory in Norman, Oklahoma, annually for 6-8 weeks during the spring severe weather season (Tarp, n.d.).

The Hurricane and Ocean Testbed (HOT) is a new testbed at the NHC. The HOT has similar goals to the HWT: testing new products that could provide solutions for common challenges in high-impact tropical/marine weather and ocean condition forecasting (AOML Communications, 2022). Researchers using the HOT aim to gain feedback and insight into how forecasters use meteorological models and tools in a realistic, operational setting. What makes the HOT unique from the HWT is its goal is to more formally host and support social science studies that focus on the societal impacts and mass communication required to help protect life and property (AOML Communications, 2022).

The HOT has both physical and virtual space for researchers and forecasters from NOAA and the NHC to collaborate on joint projects to optimize tropical analyses, forecasts, and warnings. As the HOT is new, it has fewer established methodological practices, affording the opportunity to construct innovative research designs. Beyond the affordances of the facilities at the NHC, the HOT is located in a hurricane-prone part of the country; EMs and meteorologists who routinely deal with hurricanes are closer to the testbed and are easier to access.

For this study, I conducted semi-structured interviews outside of the testbed due to constraints of the hurricane season (discussed in more detail in Section 4.1). Semi-structured interviews follow an interview protocol, with the possibility for conversation and additional probing (Cramer, 2017). EMs and meteorologists from hurricane-prone regions were invited to a virtual space to use and give feedback on the prototypes of the wind exceedance graphic. These prototypes varied in probability language, color schemes, and landmarks, as discussed earlier in this dissertation. Prototypes were designed to reflect the most common locations these user groups come from to assist the localization efforts (e.g., Gulf/Atlantic states, including Louisiana and Florida). Participants were asked quantitative items for the concepts related to the study such as spatial cognition and numeracy, in addition to demographics like age, gender, work experience, and education. Additionally, participants were asked qualitatively about what additional information they gained from the graphic and how it affected their risk perceptions. I outline the procedures I used in the interviews in Section 3.5 of this dissertation. The interviews were conducted in October and November of 2022 to discuss hurricane products during the active hurricane season when the threats were more relevant to the specialized groups.

3.3 Public Testing of the Graphic Prototypes

The same prototypes of the graphic that were tested on the EMs and meteorologists were tested on a general public sample from a hurricane-prone region, with one exception: the windspeeds were shown in miles per hour instead of knots. This change in units was recommended both by the experts in their semi-structured interviews, as well as by the NHC personnel in their informal interviews. The purpose of collecting data on the general public's response to the graphic was to test whether there were differences in the design preferences and interpretations of the wind exceedance graphic product for a non-expert group when compared to an expert user group. I outline the procedures I used in the public test in section 3.6.

I looked into different online panels to see which would be the most representative for my sample. Qualtrics uses a recruiter to draw from different panel services to compile a sample for the researcher (Peer et al., 2021). While Qualtrics panels are costlier than other comparable online panels like Amazon's MTurk, Qualtrics panels can be more representative of the U.S. population because researchers can provide demographic quotas to Qualtrics that match specific geographic areas (Smith et al., 2016; Zack et al., 2019). The U.S. population has an average age of M = 38.1 years old (median is 38.6 years old), whereas MTurk does not allow researchers to stratify their samples like Qualtrics does. As a result, MTurk examples are younger on average (M = 36.49 years old, median not provided in study) than the U.S. population (Chandler et al., 2019; O'Neill, 2022). Qualtrics panels are older on average; in a random sample of Qualtrics participants without demographic quotas, Peer et al. (2021) had over half of their sample over the age of 64 years old (Peer et al., 2021). MTurk panels were found to have only about 4.3% of participants over the age of 65, while the U.S. had 16.21% of its population in this cohort at the time the reference study was published (Chandler et al., 2019; O'Neill, 2022).

Qualtrics better mirrors the U.S. age distribution; 70% of the MTurk population is below 40 years old, whereas only 35% of the U.S. population is below that threshold (Chandler et al., 2019; Peer et al., 2021). In a study comparing MTurk to Qualtrics, Smith et al. (2016) found that the median annual household income was also much higher for the Qualtrics panel (Median = $50,000 – 59,999 per year) and closer to the U.S. population (Median = $57,617 at time of publication) than MTurk (Median = $40,000 – $49,999 per year) (Guzman, 2017). However, the race and ethnicity characteristics of the sample were less diverse than for MTurk (Smith et al., 2016).

Ultimately, I used Qualtrics because this online panel service allowed me to use demographic quotas to be representative of residents of Florida and Louisiana. To get the standard sample for a general population, I set the quotas to mirror the census demographic quotas (quotas are allowed to be plus or minus 5%, which is an expected threshold for Qualtrics) (United States Census Bureau, 2021b, 2021a). I used nonproportional sampling for the two states, meaning equal representation from Florida and Louisiana despite Florida having a larger overall population. Because I wanted to determine whether I could combine the samples from the two states, having equal sample sizes reduced the likelihood that standard inferential tests would fail to meet the underlying assumption of equivalent error distributions in the two samples. The age of participants was stratified to mirror the state averages: ages 18-34: 30%; ages 35-54: 32%; ages 55+: 38%. Race quotas matched census percentages for Louisiana and Florida: White: 75%; Black/African American: 13%; Asian or Pacific Islander: 6%; Native American/Alaskan Native or Other: 6%. Overall, the sample set quotas for 18% Hispanic and 82% Non-Hispanic participants. The last quota was gender, which was set to 50% male and 50% female. By the end of data collection, I recruited 624 participants for this public sample.

3.4 Measurement of Variables of Interest

3.4.1 Numeracy

For this study, numeracy is defined as numeric literacy, meaning one's ability to process and understand basic probability and numerical concepts (Peters et al., 2006). Numeracy levels influence the judgment/decision-making of individuals when they are presented with data and probability language (Grounds et al., 2017; Joslyn et al., 2009; Peters et al., 2006). Risk is sometimes misestimated by those with lower numeracy levels, so it is important to measure participants' numeracy to investigate the relationship between numeracy and the interpretation of the wind speed exceedance graphic (Ripberger et al., 2022).

Numeracy has been measured in other studies mostly through objective measures, like mathematical tests (Fagerlin et al., 2007). Some common mathematical tests that have been used include answering items that convert percentages to proportions and vice versa, calculating fractions, and converting probabilities to percentages and proportions (e.g., Lipkus et al., 2001; Peters et al., 2006; Schwartz et al., 1997). Mathematical testing is a direct way to assess numeracy and has been used successfully in studies of risk perception; however, many authors have found problems with using this approach (e.g., Fagerlin et al., 2007; Lipkus et al., 2001). One major issue is the additional cognitive load put onto participants when they must take a math test, which has led to lower completion rates and higher attrition rates in longitudinal studies (Fagerlin et al., 2007). Also, there have been concerns with using these tests in online experimental designs since participants could use a calculator or look up the answer in another tab without the researcher knowing, inflating their numeracy levels (Fagerlin et al., 2007). Thus, subjective scales for numeracy were developed to provide a quicker, less difficult-to-obtain measure for numeracy that still accurately reflects numeracy levels (Zikmund-fisher et al., 2007).

In this study, participants were asked to self-report their level of numeracy. The subjective numeracy scale (SNS) is an 8-item Likert scale that assesses subjective numeracy through participants' reported cognitive abilities and preference for display of numeric information (Fagerlin et al., 2007). The SNS has been well-validated through multiple, high-powered surveys of representative samples. These studies have shown that the SNS predicts objective numeracy abilities (Zikmund-fisher et al., 2007). Important for this study, the SNS has been used in the forecasting domain in a study of how EMs deal with uncertainty and probabilistic formats for forecasting in an HWT testbed experiment (Adams et al., 2017). The SNS has also been used in a risk communication study; Hess et al. (2011) determined in their study on graphical risk ladders that the self-report scale was the most appropriate for interpreting risk graphics because it relates to self-efficacy. The SNS provides information about both the ability and the self-efficacy aspects of numeracy (like confidence), and self-efficacy measures have been shown to be relevant to information-seeking behaviors, which is a common desired outcome of risk communication (Hess et al., 2011). Thus, using a subjective numeracy scale appears to be appropriate for this study as it has been found to work in both the populations and domains of this study. The scale for subjective numeracy used in this study is located in Appendix 7.1.

3.4.2 Spatial Cognition

Spatial cognition is defined as the ability to effectively use, navigate, and interpret information in a spatial representation, like a map (Tomaszewski et al., 2020). This concept has gone by different names, such as spatial thinking ability or sense of direction. Additionally, spatial cognition has been measured in two primary ways, using objective and subjective assessments. For example, spatial thinking ability was developed as an objective measure to describe a person's ability to locate, navigate, and understand the relationships between objects in space (Tomaszewski et al., 2020; Turgut, 2015). Spatial thinking ability has been measured through tests like the spatial thinking ability test (STAT), a 16-question, standardized assessment by Lee and Bednarz (2012). STAT spans eight aspects of spatial thinking abilities: (1) comprehending orientation and direction; (2) comparing map information to graphic information; (3) choosing the best location based on several spatial factors; (4) imagining a slope profile based on a topographic map; (5) correlating spatially distributed phenomena; (6) mentally visualizing 3-D images based on 2-D information; (7) overlaying and dissolving maps; and (8) comprehending geographic features represented as a point, line, or polygon (J. Lee & Bednarz, 2012). The STAT has been used mainly in the context of education, measuring students' abilities at both high school and university levels (J. Lee & Bednarz, 2012).

Despite the use of STAT across disciplines, there have been critiques of the test. In a companion publication to Lee and Bednarz (2012), who developed STAT, Bednarz and Lee (2011)[5] found little empirical support for the components of STAT. Through a principal component factor analysis of a large student sample, the authors could not confirm the reliability or the validity of the components in STAT. Even the authors noted that they were *not* suggesting that STAT was the "optimal assessment instrument for uncovering the components of spatial thinking" (Bednarz & Lee, 2011, p.105). For example, some students use verbal strategies to solve spatial problems, which cannot be captured in STAT (Bednarz & Lee, 2011). At the time of writing this dissertation, I found no other published studies about this test's validity. STAT combines items and concepts from other objective measures of spatial thinking (e.g., Gersmehl,

[5] Interestingly, the authors released a critique of STAT in 2011 through conference proceedings before the test was published in 2012 in the *Journal of Geography*.

2014; Golledge, 2002) and is still widely used despite the limitations to the test. Objective measures may not be the most appropriate test to use for spatial cognition in this study. Thus, to more robustly capture this ability, this study took a self-report approach to capturing spatial cognition, similar to the one proposed to capture numeracy.

Sense of direction has been measured subjectively through self-report items, which have shown to be a more promising approach to predicting spatial cognition (Hegarty et al., 2002, 2018). The Santa Barbara Sense of Direction (SBSOD) scale, located in the Appendix, is widely used, reliable, and has been validated empirically in other studies (Hegarty et al., 2002; Weisberg et al., 2014). The SBSOD also correlates with the ability to locate landmarks, describing not only a general sense of direction, but people's navigation ability and comfort with operating spatially (Hegarty et al., 2002). While there are some concerns around how predictive self-report scales are when compared to tests of spatial cognitive abilities (e.g., Taekeuchi, 1992; Thorndyke & Goldin, 1981), Hegarty et al. (2002) found in their multi-study experimental design that people were actually quite accurate in their assessments of their own environmental spatial abilities. This result suggests that the SBSOD (Appendix 7.2) is a useful and easy way to assess spatial cognition.

3.4.3 Experience with Hurricanes

For this study, experience is defined as the knowledge of an event or process gained from direct or indirect participation or observation (Vasileiadou & Botzen, 2014). There are different contexts in which people gain experience—work, school, home, etc. Regardless of where people gain it, experience informs how they respond to and assess stimuli (Vasileiadou & Botzen, 2014). For example, previous studies of risk perception have measured experience as an extension of the availability heuristic, finding that probability information about the occurrence of a flood was perceived as riskier when participants had prior flood experience (Keller et al., 2006). Experience has been measured for hazards like flooding and earthquakes, similarly finding that experience was associated with higher risk perceptions for the hazard (Keller et al., 2006; Siegrist & Árvai, 2020; Tanner & Doberstein, 2015). Since not all severe weather warrants the same experience, measures that are more specific to a hazard experience are better for describing the surrounding context. For example, measuring tornado experience specifically in a study on tornadoes is a better practice to describe the tornado experience than would be a general scale (Demuth, 2015). Two contexts that inform people's experience are the personal and professional contexts, which have been shown to affect disaster risk perceptions (Vasileiadou & Botzen, 2014).

Professional experience describes the events and interactions people have from their involvement with a particular job (Vasileiadou & Botzen, 2014). Those whose jobs or careers put them at greater risk because they work outdoors or have them managing a hazard in their profession will have a different level of experience with a natural disaster. For example, farmers will have different experiences with natural disasters because they may have to work outside during a natural disaster in order to protect their fields and livestock, whereas someone whose work is not necessarily affected by a hazard, like a writer or teacher, would not have to be outside during the storm. EMs or meteorologists would have an even more different experience during a hazard than the professional positions mentioned before, as they may be required to stay in a city during a natural disaster to coordinate resources and manage communication between the different sectors of their community. EMs and meteorologists, therefore, may have both personal and professional experiences in a disaster due to their professions.

Professional experience, in this study, is defined as involvement and direct knowledge of events or processes of a natural hazard (e.g., hurricane) that comes from having to address the hazard through their work. Literature in the finance domain has measured professional experience as the number of years in a given job position or the number of events one went through in that job position that can be classified as distressing (e.g., investment crashes) (Dittmar & Duchin, 2016; Faulkner, 2019).

Personal experience, though measured quite a bit in the risk and hazard scholarship, has few formal definitions. Illustratively, in one of the canonical papers on personal experience's effect on self-protective behaviors, personal experience was described as a "reasonably well-defined concept," though there was no definition of this term included in the paper (Weinstein, 1989, p.32). Hazard experience has been defined as the recency and frequency of the impacts from a hazard that the person themselves, their family, or other social networks went through (Lindell & Hwang, 2008). However, most studies that measure hazard experience do not ask about the recency or frequency specifically (Demuth, 2015). For example, in studies about severe weather, experience was measured using general, yes/no questions like "Have you ever experienced a hurricane?" and "Has someone you know ever experienced a hurricane?" (Goddard, 2017), or "Neighbors or acquaintances were harmed by a hurricane," and "The house I am living in/used to live in or my personal property was once damaged by a hurricane" (Keller et al., 2006). These items are often used as a quantitative ordinal variable to categorize levels of experience when the yes/no answers are summed to create a score.

Hurricane-specific scales have been created to measure personal experience, like the Hurricane Experience Scale (HES) (Ehrlich et al., 2010). This scale's items asked (yes/no responses, which were summed) if participants feared for their life due to the storm, suffered or had a household member or relative suffer an illness, walked in floodwaters, saw someone die, knew someone who died, or had major damage to their house. The HES was developed in a study of residents of Dade County, Florida, after Hurricane Andrew and later was validated in a study of women in southern Louisiana after Hurricane Katrina (Ehrlich et al., 2010; Norris et al., 1999). Thus, HES was determined to be an appropriate scale to use in the public prototype test (Appendix 7.4).

Qualitative studies of extreme weather have used a different approach to measure experience. Participant responses about experience can reveal more of the nuances of personal experience through the adding of descriptive adjectives (e.g., “intense”) when describing their encounters with natural hazards (Vasileiadou & Botzen, 2014). Vasileiadou and Botzen (2014) coded qualitative responses in their interviews on personal experiences with severe weather, capturing the variable of intense personal experience based on whether the individual described an intense, life-threatening experience with an extreme weather event in the past. A qualitative approach to measuring personal hazard experience was helpful in their study because the interviews allowed participants to elaborate on their different feelings they had during an extreme weather event. The way they shared their feelings and experiences helped the authors to determine the role intensity played in participants’ experiences (Vasileiadou & Botzen, 2014). Thus, this study captures personal experience both quantitatively using scales in the study of a public sample and qualitatively through open-ended questions in the interviews with the expert user group.

3.4.4 Accuracy of Interpretation

Accuracy can be defined as the correctness someone has in their calculations (Schwartz et al., 1997). This study is concerned with users’ accuracy in interpreting the wind exceedance forecast graphic, meaning how right the user is when they are asked to determine a value from the map graphic. To measure accuracy, participants were asked what the wind exceedance was at two designated locations. Their answers were compared to the correct value. Using the legend on the graphic, the answer was given an accuracy score. Participants were given a zero if they were correct, and their answer was scored one point for each category[6] the answer was away from accurate. For example, if the windspeed was in the 34-50 knot range and the participant said it was 90 knots or greater, their assessment was scored a 3 as it was 3 categories away from correct. An explanatory figure for this score is shown in Figure 14 in Appendix 7.12.

3.4.5 Confidence in Interpretation

For this study, confidence is defined as a participant's level of assuredness in their ability to accurately assess the meaning of the data in the wind exceedance graphic. Confidence has been measured in previous studies both directly and indirectly. Bisantz et al. (2005) considered high confidence to be when their participants purchased a high amount of stock after being shown a probabilistic display of stock information, where "high" was determined mathematically from the upper bounds of the average purchases made by participants in the sample. The argument for this indirect measure of confidence was that this was an impartial way to determine how confident participants felt; the researchers stated that they could objectively determine how confident participants were after exposure to the experimental stimulus based on their behaviors, e.g., on how many units they bought, rather than a self-report (Bisantz et al., 2005).

While indirect measurement is generally considered to be more objective, the subjective assessment of confidence through direct measures is considered an effective way to capture

[6] Windspeed category values and the number of categories were selected by CIRA based off recommendations from the HSU.

confidence, especially in the uncertainty visualization literature (Kinkeldey et al., 2017). Confidence is most often assessed with self-report items; it has been measured by asking participants about their confidence that their answers to a question are correct. In these cases, participants use either a Likert scale or numerical-input items where participants provide the percentage they feel confident in their answers (e.g., Ferreira & König, 2014; Kinkeldey et al., 2017; Riveiro, 2016; Roth, 2009).

For this study, I measured confidence using the self-report Likert scales. Measuring confidence this way has been done in previous studies, such as using a scale from 1 (not confident at all) to 9 (very confident) (e.g., Ruisch & Stern, 2021) or using a scale from 1 ("random guess") to 5 ("sure of answer") (e.g., Salovich et al., 2021). Roth (2009) used a five-point Likert scale asking, "On a scale of 1-5, how confident are you that your decision is correct?" (Roth, 2009). In this study, I modified the scale to reflect how confident the participant was in how accurate they were when interpreting the graphic. For example, participants were asked "How confident are you that your answer is correct?" Confidence was assessed for both the general public's and the expert user groups' (EMs and meteorologists) interpretations of the graphical information using a 5-point scale.

3.4.6 Risk Perceptions

Risk perception describes someone's assessment of their level of risk for a particular hazard (Retchless, 2014; Roth, 2009). Hurricane risk perception has been directly measured using a three-item scale created and validated by Peacock et al. (2005). In their study, Peacock et al. (2005) wanted to find out whether the location of homeowners in Florida influenced their risk perceptions for hurricanes, along with other factors such as their knowledge of hurricanes, previous hurricane experience, and socio-economic and demographic characteristics (Peacock et al., 2005). The researchers measured risk perceptions on a 5-point Likert scale (from very unlikely to very likely) with the items "How likely do you think it is that a hurricane will prevent you or members of your household from being able to go to work or go to your jobs during the next hurricane season?," "How likely do you think it is that a hurricane will disrupt your daily activities during the next hurricane season?," and "How likely do you think it is that a major hurricane will potentially damage your home during the next hurricane season?" (Peacock et al., 2005). The empirical results from validating the scale in this study suggested that the items hung together well; the inter-item correlations ranged from a low of 0.42 to a high of 0.56, with an average correlation of 0.47 and a Cronbach's alpha of 0.73 (Peacock et al., 2005). Other studies have found this measure for hurricane risk perception to be valuable, even noting years later that this scale is the "best previously used measure at the time of the work we report here" (Trumbo et al., 2014, p.1014).

However, finding the "best" way to measure risk perceptions has been a struggle for decades (Wilson et al., 2019). In their pursuit to develop a broadly applicable measure of risk perception, Wilson et al. (2019) did a comprehensive literature review to standardize how risk perception is measured across hazards and disciplines. They found that researchers used three major approaches to measure risk perception: general risk perception, general perceptions about the dimensions of risk (probability and consequence likelihood), and affect (Wilson et al., 2019). In the first approach, researchers used general items asking how risky a hazard might be, typically in a single item measure (e.g., "How risky is X?"). In the second approach, researchers used items or scales that measured the two most generally accepted dimensions of risk: the perceived probability of experiencing a risk and the severity of the consequences (Wilson et al., 2019). In the final approach, researchers used items that focused solely on the affective response to the hazard, asking typically about negative emotions (e.g., “How worried are you about X? or “How scared are you about X?”). Wilson et al. found that the most common approach used was the first approach, which measured risk perception generally using a single item; 40 percent of the papers in their review used this approach.

Despite being widely used across domains, Wilson et al. stated that a single-item, general measure for risk perception of a hazard was too broad and may not best reflect risk perceptions between hazards. Thus, they conducted a survey of 300 participants from an online panel, asking a series of risk perception items covering affect, severity, and consequences across four hazards: extreme weather events, contaminated waterways, walking late at night in a dangerous neighborhood, and eating potentially contaminated food while traveling. To increase the applicability and salience of extreme weather events, the authors asked respondents what kind of extreme weather (e.g., flooding, high-wind events, wildfires, or earthquakes) was most common in their area and then all subsequent items about extreme weather focused on that event (Wilson et al., 2019). The authors conducted a confirmatory factor analysis comparing a three-factor model, where the three factors were the three dimensions found from the literature review (affect, probability, and consequence), against a single-factor model (a general, single-item question about risk). Items in these models of risk are shown in Appendix 7.4. The results showed that the three-factor model (affect, probability, and consequence) had an acceptable fit across all four hazard domains and did not vary across hazards (Wilson et al., 2019). However, in a regression analysis, affect and consequence consistently predicted general risk perception across hazards, but not for probability (Wilson et al., 2019). Thus, it was further supported that a multidimensional measure was a better fit to measure risk perception than a single, general item.

This study used the multiple-item scale for risk perception suggested in Wilson et al., (2019) because when asking about risk perceptions of hurricanes in general, there are a variety of hazards that would be reflected more accurately with multiple items rather than a single-item question. However, when asking about the riskiness of the graphic that used one color or one overlay, this study used a single item asking, "On a scale of 1 to 5, how risky does [color scheme/overlay option] seem?" As this measure only seeks to capture the perception of one aspect of the graphic, rather than speak towards the entirety of the graphic, a single item was deemed to be appropriate.

3.4.7 Metacognitions of Vulnerability

Vulnerable communities can be defined as sectors of the population that are at higher risk of a natural hazard due to their socio-economic status (e.g., income, housing security, access to transportation, education), geography, gender, age, or health status (Bathi & Das, 2016). As the commonly used phrase in emergency management/forecasting products for a 10% probability is "reasonable worst-case scenario," there already is an implication that something bad is going to have a negative impact on the community. In this study, I wanted to know how EMs and meteorologists think about vulnerability when they are making decisions using forecast graphics. In the interviews, I asked this specialized group to identify vulnerabilities that come to mind when they think about the impacts of hurricanes and hurricane-force winds and how these thoughts about vulnerability factor into the decisions they make in their job responsibilities.

3.5 Expert Test Procedures (Interviews)

I conducted semi-structured, in-depth interviews with EMs and meteorologists. The Institutional Review Board protocol for this study was #3592, which was approved on June 21, 2022, and an amendment to update procedures was approved on October 19, 2022. Accessing this population was done through a combination of approaches, including initial networking with meteorologists at the American Meteorological Society's 49th Conference on Broadcast Meteorology/Sixth Conference on Weather Warnings and Communication (June 14 to 17, 2022, in Milwaukee, Wisconsin). Additionally, I planned to recruit participants for interviews from the participant observation at the NHC, as the NHC has connections to FEMA personnel and other state-level EMs. This recruitment was done virtually through email introductions due to constraints from the 2022 hurricane season (See Section 4.1). Additional snowball sampling was done to reach saturation for each group. The final sample size was 19 participants, with 10 meteorologists and 9 emergency managers. The following is an outline of the interview process:

1. Potential participants were recruited, recommended, or identified via in-person or virtual networking efforts (e.g., email, telephone, or video conferencing).
2. Date, time, and place (video conference link) for a one-hour interview were set.
3. At the beginning of the interview, I read the consent script to participants (located in Appendix 7.7), which explained the study, discussed their rights as a participant, confidentiality, and risk.
    a. In the consent script, I asked the participant if I could record the interview.
    b. Participants verbally agreed to both the interview and recording device.
4. I followed the interview guide (Appendix 7.8) to ask questions, while being receptive to clarifications and follow-up questions.
    a. Some examples of what participants were asked to answer included Likert scales for some of the variables of interest (e.g., numeracy, confidence, etc.) and open-ended questions. Additionally, participants were shown the different prototypes of the wind exceedance graphic and asked to reflect on the content,

design, and how they anticipated using the forecast information. Please see the interview guide in the Appendix for a more detailed breakdown of the order and content of questions.

5. At the close of the interview, participants had the opportunity to speak about anything they thought had not been covered or emphasized enough in this study.

Of the recorded content after consenting, interviews lasted on average 0:35:53 minutes (max: 0:50:48 minutes; min: 0:22:44 minutes).

3.6 Public Test Procedures (Survey)

Participants (N = 624) were recruited using Qualtrics in communities that are impacted by hurricanes (i.e., Louisiana and Florida). Prototype graphics shown to the participants were of the same state they resided in. The public test/survey took an average of 0:13:59 minutes to complete and was launched on December 5, 2022 and closed on January 4, 2023, which is at the close of the hurricane season. Conducting the survey after hurricane season ended should have limited any potential effects that an active hurricane in the area could have had on data collection.

The questions and prototypes followed the same order as the interviews, though responses were entirely quantitative. Participants answered items about general hurricane risk perception, numeracy, and spatial cognition. The group answered items about their personal experience with hurricanes with the HES, which does not capture professional experience. Participants were shown the color and texture schemes and asked to score their risk perceptions for each graphic. Then, they were asked to interpret the graphic at locations A and B indicated on the graphic (measuring accuracy; see Appendix 7.12 for a reminder on this score) and then rate their confidence in their answers. Then participants were asked to select their preferred color scheme and then they were only shown the landmark prototypes in their preferred color scheme. Full demographics were captured as well at the end of the test (e.g., asking about career and education). Metadata from Qualtrics showed that the average response time was 0:13:59 minutes (median: 0:10:46 minutes). Survey items and prototype graphics are located in the Appendix (Chapter 7).

3.7 Analytic Approach

This study follows a mixed-methods approach, meaning that as a researcher, I am seeking to answer my research questions and hypotheses through a postpositivist, quantitative stance while also believing that qualitative data and analysis will help address the research questions to a greater extent (Hitchcock & Onwuegbuzie, 2020). Collecting quantitative data using psychometrically sound quantitative instruments, such as validated scales, during qualitative interviews has been found to augment researchers' qualitative findings by affording them more ways to contextualize the data (Frels & Onwuegbuzie, 2013). Therefore, I conducted both qualitative and quantitative analyses to address the research questions and hypotheses.

To answer the research questions, the interview transcripts were thematically coded. Interviews were transcribed using Otter.ai, and transcripts were analyzed using MAXQDA. The qualitative analytic strategy for these interviews took a maximum variation approach for capturing exhaustive themes, following a quasi-grounded theory approach (Strauss & Corbin, 1998). This thematic analysis began with a directed approach (Hsieh & Shannon, 2005) by using *a priori* codes from the guiding literature in this study, such as concepts from universal design, risk perception, etc. However, as this study asked questions at different levels of generality, an open-coding, emergent process was conducted that follows the constant comparative method of qualitative analysis (Glaser & Strauss, 1967). This inductive process captures the subjects and expressions that are unique to the hurricane context, which allows for a richer understanding of the relationship between expert users (meteorologists and EMs) and their communities.

In answering the hypotheses, I checked for normality before determining what statistical tests were appropriate for the data. Due to differences in sample sizes between the large public sample and the smaller expert group, non-parametric tests were used for the expert sample and parametric tests were used in the public sample.

3.8 Reflections from the Pilot Tests of the Public Test and Expert Test

The public test/survey was piloted using a convenience sample to check that the time taken, understanding of question wording, number of successful completes, and survey logic was functioning before it went to Qualtrics for fielding. The survey was tested both for desktop and mobile phone use and no changes were made to the survey based on the piloting. Qualtrics ran a soft launch (n = 30) and showed me the data as a safeguard on the quality of the data gathered. The median time to complete the survey in the soft launch was 00:09:00 minutes and there did not appear to be issues regarding speeding.

Pilot testing of the interview guide was done prior to data collection to test the appropriateness and understandability of the questions, the length of time the interview would take, as well as the technology required to conduct and record the interviews. Using prototypes for case storms in Florida and Louisiana, I tested the interview guide with 5 individuals. These pilots were done with two meteorologists and one emergency manager who lived in Louisiana, as well as with two CSU professors who are familiar with Florida and had familiarity with hurricanes. The main lessons I gleaned from the pilots included how to navigate between interview questions more naturally, how to measure accuracy with the map more efficiently, as well as timing the length of the full interview, which includes procedural, non-recorded content as well as recorded time. The average time of the recorded content in these pilot interviews was 00:24:32 minutes. The time spent answering the quantitative items was similar for both those who answered the items remotely beforehand and those who answered the items during the interview (M = 00:04:55 minutes).

The first two pilot interviews took place on August 29 and 30, 2022, with Dr. Jeff Pierce and Dr. Bonne Ford from CSU. They have backgrounds in atmospheric science, even if their research is not focused on tropical or synoptic meteorology. Both Pierce and Ford said they were more familiar with the geography of Florida (vs. the Gulf Coast), so I used the prototypes of Hurricane Irma instead of those from Ida.

In these two pilots, I wanted to test the flow of the interviews with participants over Zoom (i.e., was there a difference in rapport if participants came having done the questionnaire ahead of time or not), as well as to make sure questionnaire completion would not prime them. I piloted with Pierce answering the quantitative items during the interview in a different window, and Ford answering them beforehand. As for whether answering the questionnaire interrupted the flow of the interview or primed the interview, Pierce said he was unbothered by taking the questionnaire during the interview in another window. He also did not appear to be primed by the items, as he did not bring up words or phrases from the items in his responses during the interview.

On September 7 and 13, 2022, Dr. Nicolas Sokol and Dr. Joseph Harris agreed to pilot the interviews with me so I could gain more knowledge of interviewing people who worked in meteorology, cartography, and emergency management in Louisiana. Sokol is a consultant in California for solar farms with private forecasting, as well as founder/CEO of a platform on how to use algae sustainably. His MS in geography was earned at Louisiana State University and his

PhD was at the University of South Carolina in geography. Harris' PhD is in geography from Louisiana State University and his MS in geosciences was from East Tennessee State University. Harris is an assistant professor of emergency management at Massachusetts Maritime Academy, and he also works with the Stephenson Disaster Management Institute at Louisiana State University. He is skilled in GIS and has worked on several hazard mitigation plans in Louisiana, Tennessee, and North Carolina. The prototypes of Hurricane Ida were used in these pilot interviews.

Sokol answered the quantitative scales during the interview in a different window, and Harris answered them beforehand. As for whether answering the questionnaire interrupted the flow of the interview, it did not appear to do so. When asked about how answering during the interview felt, Sokol said it was fine and not a problem. Harris just thought the items were interesting. Completing the questionnaire ahead of time did not appear to prime his answers during the interview because neither Sokol nor Harris brought up concepts or phrases from the items in the questionnaire in their interview responses.

After four pilots were completed, a fifth pilot interview was conducted because there were updates to how accuracy was measured. In this pilot, I tested measuring accuracy by asking the participant to identify the windspeed on a grayscale map for points A and B, which were in color in previous pilot interviews. On September 27, 2022, Holly Mallinson agreed to pilot the interview with me; Mallinson is a PhD candidate at the University of Illinois Urbana-Champaign in the Department of Atmospheric Sciences. Prior to this, she attended the University of Louisiana-Monroe for her bachelor's degree in atmospheric science while minoring in history. Her meteorological skillset and local knowledge of Louisiana made her an excellent person to test the interview. The prototypes of Hurricane Ida were used in this pilot interview.

Mallinson answered the quantitative items before the interview, and again, I did not notice a priming effect of the questionnaire on the interview. This time, I tested using the feature of Zoom annotations[7] as a tool in the interview. Unfortunately, Zoom annotations were not helpful, and I did not use them with the study participants. While this annotation tool was only tested in Mallinson's pilot, the normal functionality of Zoom for recording was practiced in each pilot. Every pilot interview served as a technological test to make sure the recording, sound quality, screen-sharing, and transcription software worked, which they did.

To summarize the main procedural changes to the interviews from the pilots, the order of some questions was changed to improve the flow of the interview, grayscale graphics were used to measure accuracy before the color schemes were introduced, and I decided to have participants answer the quantitative questionnaire ahead of the interview.

3.9 Expert User Sample

Nineteen experts were interviewed, including ten meteorologists and nine emergency managers (EMs). The length of their careers was 19.37 years on average (median: 20 years, max: 33 years, min: 3 years). The sample was predominately male (78.95%; n = 15). No one reported having a color-vision deficiency or being colorblind. Of the sample, 6 experts were from Louisiana (31.58%) and 13 were from Florida (68.42%). This unequal representation reflects the fact that Florida is more populous, has more counties, and has more NWS weather forecasting offices than Louisiana.

[7] Zoom annotations is an in-meeting product feature that allows participants to add annotations on the shared screen during video calls, including stamps, drawing, text boxes, etc.

3.9.1 Reliability and Normality of Expert Quantitative Data

Each participant answered the quantitative scale items, which were analyzed using IBM's SPSS version 26 software. Each scale had a reliable Cronbach alpha value (Table 1). To assess the normality of the data, a Shapiro-Wilk test was used as it is highly recommended for sample sizes under 30 (Ghasemi & Zahediasl, 2012). Table 1 contains both the alpha and W statistics. The distribution of numeracy, spatial cognition, and risk perception were statistically non-significant. Therefore, I can reject the alternative hypothesis and conclude that the data comes from a normal distribution, so parametric tests are appropriate.

Table 1. Reliability (alpha) and Normality (W) Statistics for Expert Sample Quantitative Data.

Variable	**Scale**	**Cronbach's α**	**Shapiro's W**	**p-value**
Numeracy	SNS (Fagerlin et al., 2007)	0.77	0.95	0.45
Spatial Cognition	SBSOD (Hegarty et al., 2002)	0.75	0.96	0.58
Risk Perception	Risk Perception (Wilson et al., 2019)	0.73	0.95	0.44

3.10 Public Sample

There were 624 participants in the public survey. I set the selection criteria for participation in Qualtrics to follow the census data recommendations in Florida and Louisiana to represent the demographic quotas for gender, age, and race/ethnicity, as discussed in section 3.3 (United States Census Bureau, 2021b, 2021a). The final sample had a relatively even gender spread (50.30% male, 48.60% female, and 1.10% non-binary/third gender or prefer to self-describe), which reflected both Florida and Louisiana. Only 40 participants (6.70%) reported having a color-vision deficiency or being colorblind, which is close to the national average (Crameri et al., 2020). Of the sample, 299 participants were from Louisiana (47.90%) and 325 were from Florida (52.10%). The average age of participants was 45.84 years old, which is similar to the age distribution in both states (SD = 17.66 years; min: 18 years old, max: 89 years old; median: 44 years old). As seen in Appendix 7.10, the race/ethnicity distribution of this sample was similar to the race/ethnicity distribution in Florida and Louisiana as well (United States Census Bureau, 2021b, 2021a). Other demographic data for the survey sample, such education level, employment status, and income are also listed in Appendix 7.10.

3.10.1 Reliability and Normality in Public Quantitative Data

The quantitative scale items were reliable in the public sample (Table 2). Histograms were run for each measured variable to see if the data appeared normal, also known as visually checking the data. Additionally, the large sample size points towards normality. If the data follow these statistical criteria, normality can be assumed (Ghasemi & Zahediasl, 2012).

Table 2. Reliability Statistics for Public Quantitative Data.

Variable	**Scale**	**Cronbach's α**
Numeracy	SNS (Fagerlin et al., 2007)	0.81
Spatial Cognition	SBSOD (Hegarty et al., 2002)	0.86
Risk Perception	Risk Perception (Wilson et al., 2019)	0.85

As this survey covered two Gulf/Atlantic states, to boost the generalizability of the findings, an independent samples t-test was conducted to test whether there were statistically significant differences between the samples from the two states so I could combine the sample. There was not a statistically significant difference in numeracy, spatial cognition, nor risk perception between the Louisiana and Florida participants (Table 3).

Table 3. Independent samples t-test values for scales in public survey data.

Variable	State	M	SD	t	df	p-value
Numeracy	Louisiana	4.01	1.04	-1.44	622	0.81
	Florida	4.13	0.99			
Spatial Cognition	Louisiana	4.50	1.05	-1.61	622	0.11
	Florida	4.63	0.95			
Risk Perception	Louisiana	3.67	0.73	1.77	622	0.08
	Florida	3.57	0.71			
Hurricane Experience	Louisiana	0.56	0.42	4.14	622	0.00**
	Florida	0.43	0.38			

**. Correlation is significant at the 0.01 level.

However, there was a statistically significant difference in the scores for hurricane experience between participants living in Louisiana and those living in Florida. Cohen's D was calculated for this analysis ($d = 0.60$), meeting the standard for a medium effect size (Cohen, 1988). Thinking about where the difference could stem from regarding hurricane experience, I checked whether participants' length of time living in these states explained this variation. A Pearson correlation coefficient was computed to determine the relationship between participants' length of time residing in their state and their hurricane experience score. There was a statistically significant, weak positive relationship between length of time in state and hurricane experience, [$r(610) = 0.12$, $p < 0.01$]. The average length of time in Louisiana was 20.17 years (SD_{LA} = 17.37 years). This was higher than length of time in Florida (M_{FL} = 11.28 years, SD_{FL} = 12.51 years). Over half of the participants from Florida had lived in the state for 10 years or fewer, whereas those in Louisiana had a broader distribution of time in the state (Figure 7).

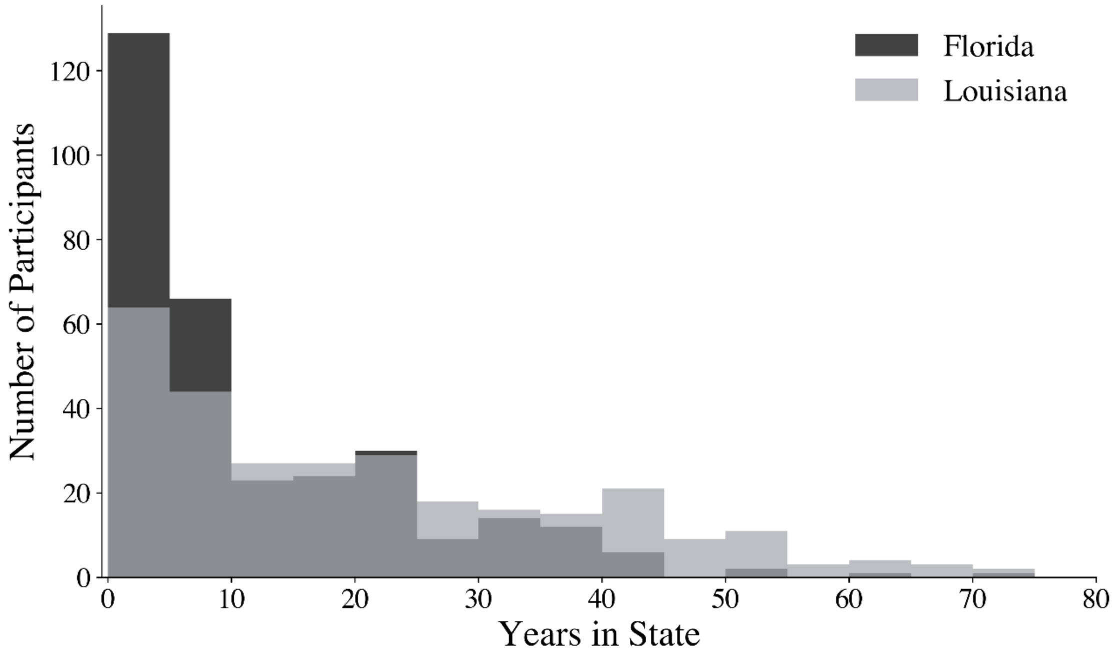


Figure 7. Overlapping histogram of number of years participants report living in their state

Age could also explain the difference in hurricane experience. The average age for Louisiana participants was older than the average age for Florida participants (M_{LA} = 48.4 years, SD_{LA} = 17.19 years, M_{FL} = 43.09 years, SD_{FL} = 17.66 years). There was a statistically significant difference in age between the participants from Louisiana and Florida $t(622)= 0.37$, $p < 0.01$. Cohen's D was calculated for this analysis ($d = -0.33$), meeting the standard for a small effect size (Cohen, 1988) However, statistics from the U.S. census show that the population is older in Florida than Louisiana (United States Census Bureau, 2021b, 2021a). Despite this difference in age or state residence, I argue that the effect size of these variables is small and negligible, so the analyses were not broken up by state residency. For the rest of this dissertation, when discussing the public sample, the groups are combined.

CHAPTER 4. RESULTS

4.1 Update to Methodological Approach

The 2022 hurricane season was unusual for its calmness. What should have been a very active season was the least active hurricane season in 30 years, and it was the first time in 40 years that no named storms formed from July 3 to August 22 in the Atlantic Basin (Erdman, 2022). This below-average hurricane activity carried into September as well. Unfortunately for my research plan, Hurricane Ian (National Environmental Satellite and Data Information Services, 2022), a category 4 hurricane that struck Florida on September 28, was the only large hurricane of the season, and this hurricane was deemed to have too many local impacts near the National Hurricane Center for me to travel to the center to observe briefings. Instead, I began conducting interviews with emergency managers and meteorologists in October 2022. I moved forward with the interviews because they were not contingent on the findings from the observations. However, despite the change in timing for the interviews, the survey was still carried out at the planned time in November 2022.

4.2 Findings Related to Hurricane Experience

4.2.1 Hurricane Experience and Risk Perceptions in Expert Users

To answer RQ1a, which asks how experience with hurricanes affects risk perceptions in expert users (EMs and meteorologists), the interview responses about experience and risk perceptions were analyzed qualitatively. When expert users were asked about the amount of experience they had with hurricanes, they often (57.89%) answered with a retort such as “Probably too much” (Expert User 2) or “I’m way too experienced I think, more experienced than I want to be” (Expert User 13). Every participant mentioned the major hurricanes they had worked in their careers by name, year, and location, except for one participant who had begun their career within the last three years. Though the question asked participants to reflect on their personal and professional experience (in that order), only three participants (15.79% of the sample) answered with their personal experience first. This pattern may be due to the occupational nature of the interview, as participants knew they were recruited to participate due to their career.

Participants brought up the frequency of hurricanes per year as a way to explain how often they gained experience with major and non-major hurricanes.[8] Each experience informs participants' frames of reference for hurricane threats over time. As Expert User 8 explained, "One of my earliest memories is with Hurricane Andrew and I want to say that was 1992. That was my first real experience with hurricanes. And, you know, it's ebbed and flowed a little bit every year since then." Participants compared different hurricane seasons to provide context for how they were describing the experience. As it was topical at the time of the interviews, participants discussed the 2022 season compared to more active years, such as 2020 (e.g., Expert Users 2, 5, 9). Notably at the time of these interviews, Hurricane Ian had hit Florida, but otherwise it was a relatively inactive hurricane season. Many participants (e.g., Expert Users 1, 2, 3, 8, 15, 19) discussed how they generally have quite a few hurricanes per season, but that this year in particular was unusual and calm. 2022 also varied in locations of hurricane landfalls from previous years. For example, Louisiana experienced major hurricane landfalls in both 2020 and 2021, with 2020 being exceptionally busy for the Gulf states. Participants from Louisiana who

[8] A major hurricane is one that is classified as category 3 or higher. Category 1 and 2 hurricanes cannot be classified as major and there is no such thing as a "minor" hurricane (National Hurricane Center, n.d.).

were working in-state for these last three years might have had a different perspective on experience than those in Florida.

While 2022 was brought up a lot because it was the current season and 2020 was the most active hurricane season on record, the Atlantic hurricane season has actually been classified as highly active since 1995 (National Oceanic and Atmospheric Administration, 2019, 2020). The previously most active Atlantic hurricane season was in 2005, with other near-record years within the last 10 years (National Oceanic and Atmospheric Administration, 2019). Most of the experts I interviewed (89.47%) had been working in forecasting or emergency management for 10 years or more, throughout these highly active years. To convey the frequency of their experiences, participants chronologically listed the storms they worked, such as:

> New York... I was there for Irene [2011] and Sandy [2012]. And then I was in Huntsville, and we had to deal with Irma [2017] remnants. And then I was in Tallahassee, and we dealt with Michael [2018]. And then I got here in 2020. So that's when we had Sally, Laura, Delta, Zeta. And then we had Ida [2021] last year as well (Expert User 2).

When discussing their personal experiences with hurricanes, participants mainly referred to personal experiences with property damage, such as:

> I mean, we're probably talking about in terms of like my house being smashed up and having to have insurance claims on it. I had that twice. 2005 with Hurricane Rita and then, and 2020 with Hurricane Laura. And you can combine it probably with the second one that hit us in 2020 as well [Delta] (Expert User 15).

Though in their quantitative ratings of risk perception for hurricanes, the expert users were only slightly above the middle-point (M = 3.67, SD = 0.51; 5-point Likert scale), they all recognized that there were risks with every hurricane experience. While the faster windspeeds were an important factor in the severity of the hurricane, other impacts such as intensification and time spent inland versus over the ocean influenced how risky they expected a hurricane to be for their areas (e.g., Expert Users 1, 13, 14, 16). In describing their experiences with hurricane hazards, such as wind and flooding, the participants' response themes matched constructs within risk perception, including affect (i.e., concern/emotion) and consequences (i.e., severity). Examples are described in further detail in the following sections.

4.2.1.1. Hurricane Experience and Affect

With the discussion of each storm, some participants responded matter-of-factly, while others responded humorously (e.g., Expert Users 2, 6, 13, 18). For example, Expert User 13 teased "I wish there were fewer storms, but they don't seem to care." Though told as a joke, by anthropomorphizing the storm, the expert's tone pointed towards more serious affective responses to hurricanes, like concern or worry.

Participants mentioned the emotional toll that experiencing hurricanes has had on them. For example, Expert User 2 described, "So [I've] kind of been around the block, unfortunately." Participants mentioned emotional tolls from major and non-major hurricanes, for example, "[Non-major hurricanes] have been less impactful personally, but stressful, and a lot of work professionally" (Expert User 12). Stress impacts emergency responders, like EMs and meteorologists, both personally and professionally, and it adds up as the hurricane season stretches on (Bolton et al., 2018). Additionally, affective responses, such as worry, were mentioned directly by participants. Like Expert User 3 described regarding the impacts in their area, "Here in Louisiana, we have a lot to worry about." Feelings of stress or worry are expected in disasters, and the expression of these feelings is apparent in both the personal and professional experiences of the expert users.

4.2.1.2. Hurricane Experience and Consequences/Severity of the Threat

Participants brought up hurricane experiences that greatly impacted them, as well as hurricane experiences that were not as severe. For example:

> I mean, I had a few, a couple of near misses…so I just remember those as a kid. But my first direct hit, you know, being directly impacted, was Hurricane Andrew and there have been several others since then. So, you know, both personally and professionally, I've dealt with a lot of hurricanes (Expert User 16).

Often, participants brought up the severity of named storms they had experienced by mentioning their category and recovery time. This information was provided for Hurricanes Andrew (e.g., Expert Users 8, 10, 11, 16), Ida (Expert Users e.g., 2, 8, 15, 19), Irma (e.g., Expert Users 2, 5, 7, 12, 15), and Laura (e.g., Expert Users 2, 3, 8, 15). Participants explained their experiences with deploying aid when tropical storms became classified as hurricanes, such as authorizing emergency responder overtime pay, opening shelters, establishing evacuation routes, or implementing emergency operations communications (e.g., Expert Users 3, 5, 8, 12, 17).

4.2.2 Hurricane Experience and Risk Perceptions in the General Public

RQ1b asked how does experience with hurricanes affect risk perceptions in the public? Measured with the hurricane experience score (HES), the public had little hurricane experience. Also shown in Figure 8, the sample's HES had a mean of 0.49, SD = 0.41, a median of 0.40, and a range of 3. Seeing as a majority of the scores were between zero and one on average, I looked within the items in the score to see if there was a most common experience within the group. The most common experience (checked "yes" in 55% of the scores) was if the participant or anyone in their household ever had damage to or loss of property because of a hurricane. The least common experience (checked “no” in 89.40% of the scores) was if the participant or anyone in their household had ever been injured, including loss of life, due to a hurricane.

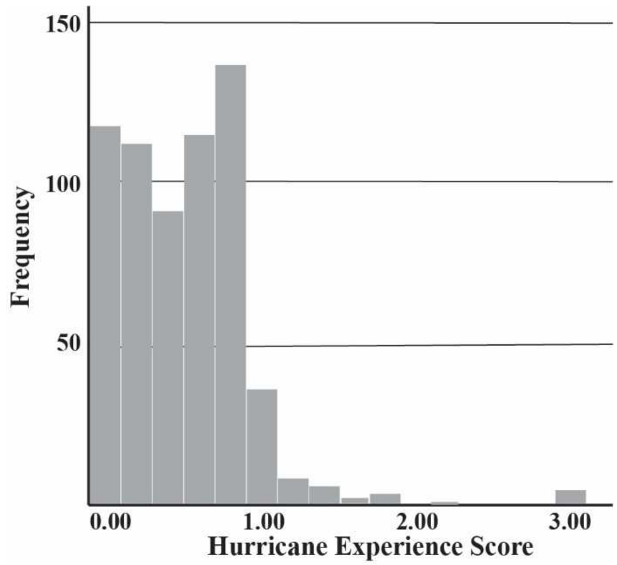


Figure 8. Histogram of hurricane experience score in public sample

Their risk perception of hurricanes was above the midpoint as well; M = 3.61, SD = 0.72. Linear regression analysis was used to test whether hurricane experience predicted participants' risk perceptions (RQ1b). The results of the regression indicated the predictor (experience) explained 7.4% of the variance in the outcome (risk perception); $R^2 = 0.07$, $F(1, 623) = 49.49$, $p < 0.01$. Hurricane experience was a statistically significant predictor of risk perception, $\beta = 0.48$, $t(622) = 7.04$, $p < 0.01$.

4.2.3 Hurricane Experience and Accuracy of Interpretation in Experts and the Public

H1 predicted that regardless of group, more hurricane experience will lead to more accurate interpretations of the wind exceedance graphic. As described earlier, the expert group has a lot of experience with hurricanes. In the expert user group, both qualitive and quantitative data was gathered to measure accuracy.

The expert user group was extremely accurate in reading the wind exceedance graphic, using the legend for windspeeds shown in the graphic. The average accuracy score was 0.08, SD = 0.25 (range: 1); the closer to 0, the more accurate. Notably, some participants said that their career training made it easier for them to accurately use the wind exceedance graphic when compared to a member of the public; for example, “I mean, it's easy to me, it's easy to interpret and discern” (Expert User 5). Participants reflected on how their high amount of experience with hurricanes gave them practice with using and deciphering maps (e.g., Expert Users 5, 11, 18). In addition to their high quantitative accuracy scores, qualitatively, experts' reflections point towards a positive relationship between experience and accuracy.

In the survey data from the public sample, a Pearson correlation coefficient was computed to determine the relationship between accuracy and experience. There was not a statistically significant relationship between accuracy ($M = 0.66$, $SD = 0.60$; range: 2.5) and experience ($M = 0.49$, $SD = 0.41$), [$r(624) = 0.03$, $p = 0.42$]. Though non-significant, the direction of the correlation was positive. Overall, H1 was only partially supported as a positive relationship between experience and accuracy could be deciphered in the expert sample but could not be confirmed statistically in the public sample.

H2 predicted that expert users (EMs and meteorologists) would be more accurate in interpreting the wind exceedance graphic than would a public sample. On average, the accuracy scores for the expert user group were closer to zero (more accurate) than the public sample, [$M_{expert} = 0.08$, $SD_{expert} = 0.25$, $Range_{expert}$: = 1; $M_{public} = 0.66$, $SD_{public} = 0.60$; $Range_{public} = 2.5$). As expected, the range of accuracy scores was also larger in the public sample than the expert group (Figure 9). While there were two outliers in the expert group, these experts have many years of experience interpreting these types of graphics and they said that their answer may be incorrect due to the difficulty they had with the grayscale as they were answering. Not only were the experts more accurate on average, but they also had a tighter distribution of accuracy scores.

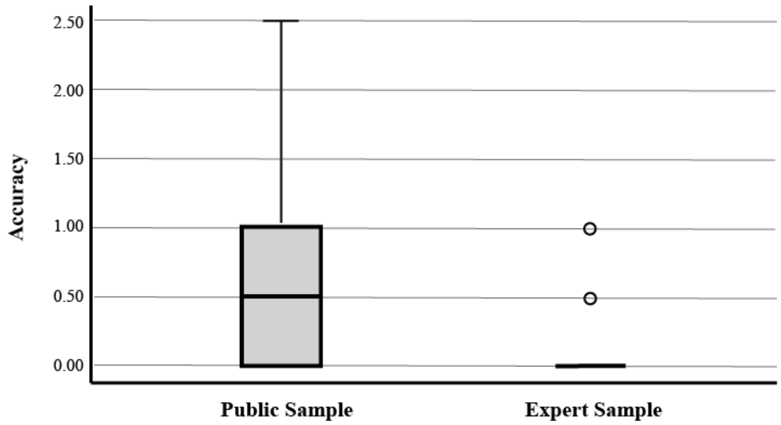


Figure 9. Boxplots of accuracy scores for the wind exceedance graphic for public and expert samples

A Mann-Whitney U test was performed to evaluate whether accuracy differed by group (expert or public). The results indicated that the expert group was significantly more accurate in interpreting the wind exceedance graphic difference than the general public group, $z = -4.53$, $p < 0.01$. Therefore, there was support for H2.

4.3 Findings Related to Confidence in Using Wind Exceedance Graphic

The third hypothesis (H3) predicted that expert users (EMs and meteorologists) will feel more confident interpreting the wind exceedance graphic than will the general public. On average, expert users were confident in their interpretation of the wind exceedance product ($M = 4.11$, $SD = 1.07$). The public was also confident in their interpretation of the wind exceedance product ($M = 4.09$, $SD = 0.99$). To test H3, I ran a Mann-Whitney U test to see if there was a statistically significant difference in confidence between the expert and the public samples. There was not a statistically significant difference in confidence between groups, $z = -0.10$, $p = 0.92$. Therefore, H3 was unsupported.

4.3.1 Relationship between Numeracy and Confidence in Wind Exceedance Interpretation for Experts and the Public

Numeracy and confidence in interpreting the wind exceedance graphic were predicted to have a positive relationship (H4). Descriptive statistics for each variable are in Table 4. The higher average numeracy score reflects the specialized training that EMs and meteorologists have, compared to a public sample. In the expert sample, a Spearman's rank correlation was computed to assess the relationship between numeracy and confidence. There was a positive, though statistically non-significant correlation between the two variables, $r(17) = 0.16$, $p = 0.52$.

Table 4. Descriptive statistics for numeracy [6-point] and confidence [5-point] measures.

Variable	Sample	M	SD
Numeracy	Expert	5.05	0.64
	Public	4.08	1.02
Confidence	Expert	4.11	1.07
	Public	4.09	0.99

A Pearson correlation coefficient was computed to test this prediction in the public sample. There was a moderate relationship between numeracy and confidence; $r(624) = 0.32$, $p < 0.01$. The expert group is a more homogenous group than the public in their mathematical training, for example, which may explain the difference in correlations. Overall, H4 was partially supported; numeracy and confidence in using the wind exceedance graphic had a positive relationship, which was statistically significant in the public sample.

4.4 Findings Related to Color Scheme for the Wind Exceedance Graphic

4.4.1 Expert User Color Preferences

RQ2a asks what are the expert user group's preferences for the color scheme and texture in the wind exceedance graphic and why did they have these preferences? Sixteen participants preferred the yellow-to-red color scheme (84.21%), followed by three who preferred the reds-only color scheme (5.79%); no one preferred the texture scheme. An excellent example of an immediate reaction to seeing the texture scheme:

> EW…Oh, I hate that one. There's never going to be one I hate more than this one…Oh, this one just screams at me. I just can't even find the county boundaries and the state boundaries are a little hard to find. The dots are distracting. I don't even hardly know what it means (Expert User 4).

The dislike of the texture scheme was also stated more tactfully:

> I like the idea of this dot one. I think this is very difficult to read, though. Because you can't tell exactly where all the dots are in some of the categories. But I like the idea of the dots, like what you're trying to do with it (Expert User 2).

Participants noted that the texture scheme had drawbacks. For example, using the pattern instead of color seemed old-fashioned, reminding them of when maps and graphics were printed solely in black-and-white due to printer cost limitations (e.g., Expert User 10). Additionally, participants remarked how the density of the pattern made the map feel cluttered (e.g., Expert Users 1, 13). In addition to the visual overstimulation from the texture, when participants were asked to describe what they were seeing further, they brought up that they were had trouble understanding the graphic quickly or communicating where they were referencing in the map (e.g., Expert Users 3, 8). Expert users need to communicate efficiently in stakeholder briefings; participants' difficulty with communicating density changes rather than color changes highlighted a major issue with the texture scheme. While the texture scheme was a good attempt at a colorblind-friendly design, it was overwhelmingly unliked by the experts.

Moving to the reds-only color scheme, three experts preferred this version of the graphic. One reason that they preferred this color scheme was that the type of hazard appeared to be the same throughout the graphic because the color was the same, meaning that the graphic was always showing differences in windspeed because there was only red (Expert Users 1 and 4). Being in the same tone category was seen as the simplest to understand and explain to their partners during briefings (Expert Users 4 and 7). Despite these reasons, the main reason that participants said they did not like the reds-only option was because it did not adequately convey the risk of the hazard.

The general impression of the reds-only color scheme was that the scheme seemed less risky than the yellow-to-red scheme (e.g., Expert Users 3, 6, 9, 12). While red is associated with danger, the lack of other colors to compare the red shades against appears to lessen the perceived risk of the hurricane represented in the graphic. Only including various shades of red was seen as potentially being less informative (Expert User 9). While yellow typically signals a lesser threat than does red, having yellow as a comparison point to red made the yellow-to-red scheme feel more severe. As Participant 13 elaborated, "With the darker reds…[there's] obviously more of an urgency, more of a serious nature to it….it stands out to where I think that that could be effective."

Notably, the yellow-to-red scheme was preferred by most of the experts. Their cited reasons for this preference was a clearer distinction between categories because of the higher visual contrast (e.g., Expert Users 1, 3, 6, 9, 10, 14, 17), the similarity to other hurricane graphics (e.g., Expert Users 5, 8, 15, 16), and their perception of being able to more quickly understand and share the information from the yellow-to-red graphic versus the other options (e.g., Expert Users 3, 11). Both meteorologists and emergency managers relayed how important it was for them to effectively craft messages for their stakeholders about the hurricane threats, and the graphic's color scheme impacted this task. For example, Expert User 5 stated, "From a messaging standpoint, I personally like this one because there's a more clear delineation of where some of the more vulnerable areas are." Using map graphics to make communication more effective and messaging easier between experts and the public is a key theme that comes up throughout the interview data.

4.4.2 General Public Color Preferences

RQ2b asks what the general public's color scheme and texture preferences were for the wind exceedance graphic. Like in the expert sample, the public overwhelmingly preferred the yellow-to-red color scheme (n = 419; 67.15%). The reds-only color scheme was the next preferred (n = 141; 22.60%), with the texture scheme being the least preferred scheme overall (n = 64; 10.26%). Figure 10 shows the distribution of color scheme preferences for both samples.

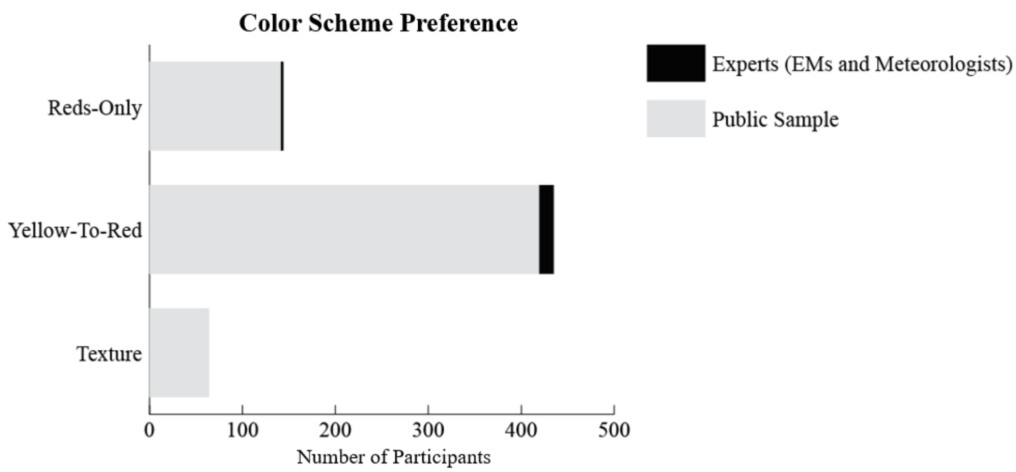


Figure 10. Stacked bar chart of color scheme preference by group

4.4.3 Effect of Color Scheme on Perceived Riskiness of Graphic

H5 predicted that regardless of group, users will perceive the reds-only color scheme as riskier than the yellow-to-red color scheme. Counter to my prediction, the expert users said that the yellow-to-red scheme felt riskier than the reds-only color scheme, due to the seemingly more drastic color contrast between windspeed categories. Expert User 17 stated:

> Yes, so you know, there because the red is so much darker than the yellow, it obviously grabs my attention more. Makes me realize that, okay, this is really important—this is a hazardous condition—which then forces you to look at your chart and see 'where are we?', as opposed to the lighter softer color, where there's less threat (Expert User 17).

This group preferred the yellow-to-red color scheme, which they felt conveyed the most risk of all of color scheme options, because of its ability to stoke an appropriate affective response from their audience and communicate the severity of the hazard. For example, "The other one [reds-only] just looks more pleasant…. I mean with [yellow-to-red], you're going to caution to more caution, whereas the other one just looks kind of relaxing and soothing" (Expert User 12). The pleasantness of the lighter red, which was sometimes described as the color pink by the participants, countered my prediction that a solely red scheme would invoke a feeling of urgency or risk. Expert User 8 summarized this point well:

> What's interesting about the yellow, the scale invokes caution to me. I don't know whether it's trained or learned. What I like about this is the fact that the red makes a nice, a nice, cool pink, kind of a mellow tone. And in my business, I sometimes need to really communicate the threat of tropical storm impacts as being somewhat serious…. So I like the yellow in the sense that it's throwing caution out here more than the muted pink or red shading of the previous graphic, which sometimes may say, 'Hey, that's a nice, warm, fuzzy color. I'm not in danger.' This [version] to me speaks at 'Hey, we're not talking about anything but various risk and danger here.' So, I say this one [yellow-to-red] stands out the most to me because it's already at a heightened state. And the red still is the ultimate, you know, highest threat there (Expert User 8).

The public scored the riskiness of the reds-only color scheme graphic higher ($M = 3.98$, $SD = 1.09$) than the yellow-to-red color scheme graphic ($M = 3.67$, $SD = 1.08$). To test H5 in the public sample, I ran a Wilcoxon matched pairs signed rank test to determine whether there was a difference in the scored riskiness of the reds-only and yellow-to-red color schemes. There was a significant difference in perceived riskiness of the color schemes, $z = -6.48$, $p < 0.01$. The results indicate that the reds-only color scheme was perceived as significantly riskier than the yellow-to-red color scheme. Thus, H5 was partially supported; while the experts perceived the yellow-to-red scheme as riskier than reds-only as expected, the public did not.

H6 predicted that regardless of group, there will be a relationship between color scheme preference and perceived riskiness of the color scheme[9]. Qualitatively, the yellow-to-red color scheme was overwhelmingly preferred by the expert user group (16 of 19 participants), and it was perceived as showing more risk than the other color schemes, as discussed in the previous paragraphs. In the public group, as a reminder, perceived riskiness was measured quantitatively via a single item that asked specifically about the riskiness of the color in the graphic. While general risk perception is better measured using multiple items in a scale, single-item measures have been used in much of the risk scholarship for measuring direct aspects (Wilson et al., 2019).

Overall, I found that preference did not follow perceived riskiness. In the public sample, participants who preferred the reds-only color scheme rated that scheme as the riskiest, followed by the yellow-to-red color scheme, and then the texture scheme (Table 5). Those who preferred the yellow-to-red color scheme rated the reds-only color scheme as the riskiest on average as well, also followed by the yellow-to-red color scheme, and then the texture scheme. Lastly, those

[9] Remember that all color scheme options are included in H6, whereas only the yellow-to-red and reds-only schemes are in H5.

who preferred the texture color scheme followed the same pattern as the previous two groups, rating the reds-only color scheme as the riskiest on average, followed by the yellow-to-red color scheme, and then the texture scheme.

Table 5. Descriptive statistics for riskiness [5-poimt scale] of color schemes, by preference group in public sample.

Preference Group	Variable	M	SD	n
Reds-Only	Perceived Riskiness of Reds-Only Scheme	4.14	1.02	141
Yellow-to-Red		3.60	0.98	419
Texture		3.47	1.22	64
Reds-Only	Perceived Riskiness of Yellow-to-Red Scheme	3.91	1.07	141
Yellow-to-Red		3.72	1.11	419
Texture		3.05	1.23	64
Reds-Only	Perceived Riskiness of Texture Scheme	4.06	1.27	141
Yellow-to-Red		3.50	1.08	419
Texture		3.25	1.25	64

One-way ANOVA tests were run to assess whether there were differences between the perceived riskiness of the graphics due to color scheme preference. There was no statistically significant difference among color scheme preference groups on rating the riskiness of the reds-only scheme ($F(2,621) = 2.46$, $p = 0.09$) or for rating the riskiness of the yellow-to-red scheme ($F(2,621) = 1.55$, $p = 0.21$). However, there was a statistically significant difference between preference groups for rating the riskiness of the texture scheme ($F(2,27) = 6.40$, $p < 0.01$).

To investigate this difference in perceived riskiness of the texture scheme, I ran the Tukey-HSD post hoc test. I used the Tukey-Kramer modification of the Tukey post hoc test to investigate where the difference lay between groups because of the difference in sizes among color preference groups (S. Lee & Lee, 2018). The rated riskiness of the texture graphic was significantly lower for the yellow-to-red color scheme preference group (mean difference -0.42, $p < 0.01$) and significantly higher for the reds-only color scheme preference group (mean difference 0.42, $p < 0.01$) compared to the texture color scheme preference group (Figure 11).

Overall, a consistent relationship pattern could not be determined between color scheme preference and perceived riskiness of the color scheme. Consequently, H6's prediction was unsupported.

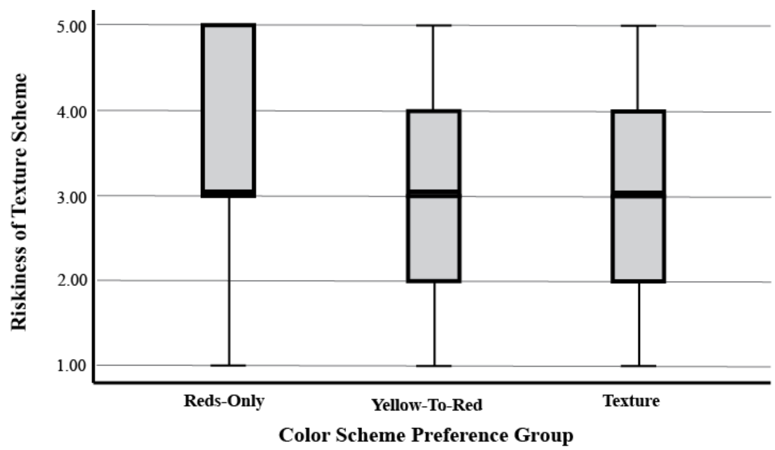


Figure 11. Boxplots comparing differences in perceived riskiness of the texture scheme, by color scheme preference group

4.5 Findings Related to Overlays in the Wind Exceedance Graphic

4.5.1 Expert User Overlay Preferences

The third research question (RQ3) asked are cities or roads preferred as landmarks for localizing the wind exceedance graphic? In the expert user group, the overlay preference was more spread out than for the color schemes: 9 (47.37%) preferred the city landmarks, 8 (42.11%) preferred the interstate/highway overlay, and 2 (10.53%) preferred neither overlay. Further description of each overlay preference is included in the following sections.

4.5.1.1. City Landmarks

Of the reasons cited for preferring city landmarks, expert users said they thought more people would know their location in relation to the cities (e.g., Expert Users 2, 5, 16) and that the city landmarks bring attention to the population centers more (e.g., Expert Users 4, 7, 18). For example, Expert User 7 explained their preference for cities:

The reason I like having cities is just because it gets those major population centers that typically are where you have ports or airports and people are just in generally familiar with them, whether it's partners or meteorologists or even members of the public. So, you know, I think it provides a good frame of reference to start with for most people (Expert User 7).

Additionally, the expert users feared that the public, which they anticipated had lower geographic knowledge, would misinterpret their location without landmarks to orient them (e.g., Expert Users 4, 9, 13). For example, "Well, they're seeing familiar names, you know, like, that makes them feel comfortable... Tampa, Orlando, Miami... Here you go—everything's defined, there's no misinterpretation" (Expert User 4).

Experts were concerned about how many cities were displayed and making the graphic feel cluttered (e.g., Expert Users 13, 16). They said that in NWS briefings, the first question that gets asked from the public is "What about my town?" (Expert Users 2, 3, 5, 16). This question is common throughout the National Weather Service for all hazards (Figure 12).

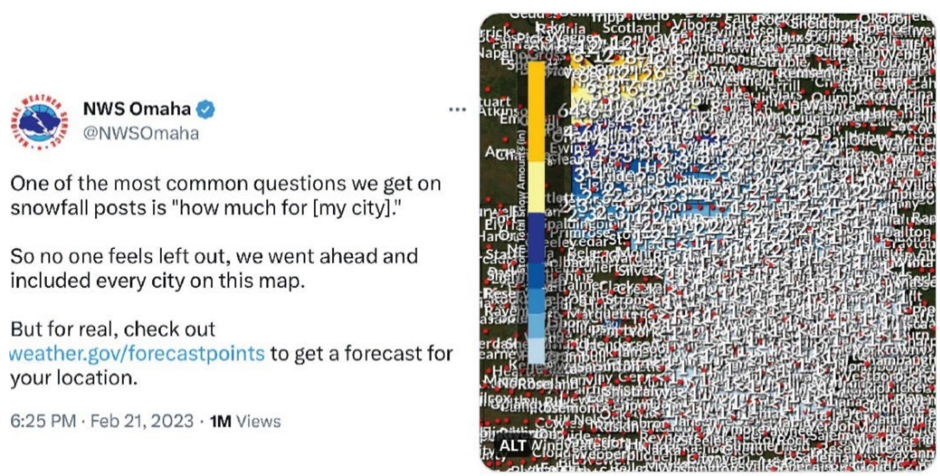


Figure 12. Tweet from @NWSOmaha about how often they get questions about unique cities

While not all towns and cities can be on a graphic, because they would crowd out the displayed data (Expert User 16), placing some cities and towns on the graphic would help people orient themselves more easily within the map.

4.5.1.2. Interstates/Highways

Of the cited reasons for why the interstates/highways overlay was preferred, the responses included ease of briefing their partners (e.g., Expert Users 2, 3, 4, 5, 6), a way to highlight impending evacuation routes more specifically (e.g., Expert Users 2, 7, 8, 9, 11, 17, 18, 19), as well as ease for the public to orient themselves on the map (e.g., Expert Users 5, 8, 9, 13). Being able to communicate the hurricane threats more easily and quickly in their briefings was important to the expert user group. Including the interstates provided more frames of reference to quickly convey where the risks were. For example:

> So I like it because from a communication standpoint, I can use simple terms like 'south of Alligator Alley and I-75' or 'south of Interstate 4.' It's a nice short clause that I can say in the briefing that might register with my partners (Expert User 6).

Additionally, as establishing evacuation orders is one major outcome of weather briefings, a graphic with interstates/highways directly on the map could be very useful for emergency managers in particular. Expert User 8 described, "From an evacuation standpoint, that makes a lot more of a noticeable sense. You know, when you're highlighting major roadways, thoroughfares, things like that, evacuation routes, you know, [the overlay] kind of helps."

Overall, the goal of using these map graphics is to help people understand the threats in their area. As such, adding interstate overlays can help residents that navigate the interstates frequently, as well as tourists who may only know the routes they took to get to their vacation (Expert Users 9, 12). Additionally, as Expert User 5 states, "We still have some folks that don't know what county they live in. But generally, at least they can know whether they're close to the interstate or not.” However, these prototype graphics did not include interstate names or numbers explicitly on the graphic. The experts, having lived in their areas for a long time, did not need the names or numbers to identify specific interstates. However, for transient populations and tourists, who may be less familiar with the road network, having this extra detail may assist them with localization and should be considered in future graphics.

Additionally, though this was not a research question, 10 participants (52.63%) asked if county boundaries could be an overlay for the graphic. For example, “I would say [the usefulness of the overlay] would depend on the audience…but if we were drilled in where we can see county boundaries, that might make a difference” (Expert User 10). County outlines form the warning-forecast areas that NWS offices oversee. Also, EMs are responsible for their counties, so they interact with these boundaries frequently. For example, “We're kind of right on the boundary…So it’s very difficult to tell [use the graphic] without a county boundary” (Expert User 14). Future iterations of the wind exceedance graphic should include county boundaries as they are highly desired by these expert user groups.

4.5.2 Public Sample Overlay Preferences

The public preferred the city landmarks overlay (n = 265; 42.47%), followed by the interstate/highway overlay (n = 136; 21.79%), and the “no overlay” option was third (n = 123; 19.71%). In addition, 100 participants (16.03%) stated that they equally preferred the two overlay options. The distribution of overlay preferences for the public sample is shown in Figure 13, including the preference of the expert group.

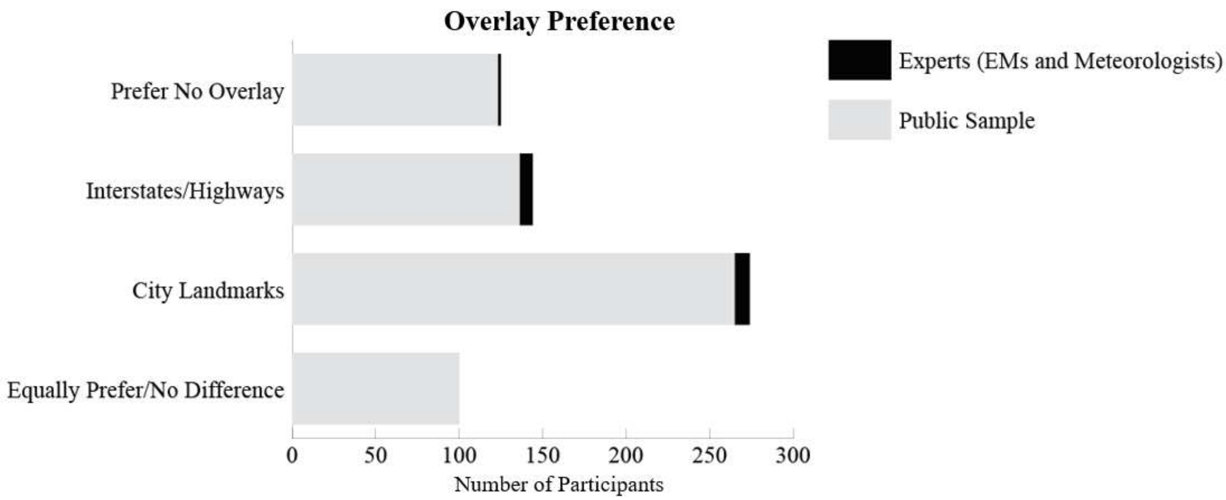


Figure 13. Stacked bar chart of overlay preference, showing city landmarks was largely preferred by the public and the experts were about equal in preference between interstates/highways and city landmarks.

4.5.3 Perceived Riskiness of Wind Exceedance Graphic with Overlays

H7 predicted that regardless of group, users will have higher perceived riskiness of the wind exceedance graphic with landmarks than for the graphic without landmarks. There were mixed perceptions from the expert users about the inclusion of overlays. For some, the risk increased when identifiable landmarks were included because they pointed out the location of forecasted threats (e.g., Expert Users 1, 9, 12). Participants talked about increased risk in the graphic for both the cities and the interstates/highways overlays. Expert User 8 stated, "Because you're highlighting major metropolitan areas, that does play into an increased sense of risk, because you're highlighting not only our capital, but also the largest city in the state as well."

In the interviews, the participants reflected that because they have a deep understanding of their areas, there may be an influence on how they perceive the riskiness of the graphics with overlays. For example, some participants compared the overlays against each other:

> I don't know if 'riskier' is the right word. You know, I don't know, I think if [prototype with interstate overlay] was the only option…I think I personally liked the cities, but I

> feel like it might be riskier if you're thinking about other parts of the country, where the roads would help those other parts of the country better (Expert User 2).

Expert User 2 felt that cities would be easier for themself than other people since they knew where the urban areas were in their state. Of note, Expert 2 is from Louisiana which has fewer major cities than other states, like Florida; they reflected that the map graphic may appear more blank in their state or other less populous Gulf states. Fewer landmarks would look as if they graphic has less detail, and it would be less useful. The usefulness of detail is elaborated on in the discussion. Also, showing that experts think about how people use maps is evidence of the way experts such as forecasters and EMs think about their constituents.

Participants said personalizing the risk with interstates as well as cities could better convey risk because traveling and commuting are frequent behaviors within populations (Expert Users 3, 5, 9). For example, "I mean, I like having the geo-reference locations on there. And I really think that helps folks interpret their localized risk" (Expert User 5). Having both overlays (cities and interstates/highways) should be considered, because the experts talked about how including both on the map could benefit the different users of the map graphic. In summary for H6, including these city and interstate/highway overlays prompted the expert users to think more about the risk of the affected communities and their potential audiences when briefing than the versions without overlays.

To test the prediction that there would be a difference in perceived riskiness of the graphic overlays (again, measured using a single item in the public sample (H7)), a Wilcoxon matched pairs signed rank test was conducted (Table 6).

Table 6. Wilcoxon Signed Ranks test for perceived riskiness [5-point scale] of overlays in public sample.

	M	**SD**	**Roads/Interstates Overlay**	**City Landmarks**	**No Overlay**

Roads/Interstates Overlay	3.76	1.03	-	-1.33	-4.87**
City Landmarks	3.75	1.05	-0.13	-	-5.10**
No Overlay	3.60	0.83	-4.87**	-5.10**	-

**. Significant at the 0.01 level.

Results of that analysis indicated that there was a significant difference in means of the perceived riskiness of the overlays between the city landmarks ($Z = -5.10$) and the interstates ($Z = -4.87$) with no overlay ($M = 3.60$). There was no statistically significant difference between the interstates and city overlays ($Z = -0.13$, $p = 0.90$). Participants viewed the versions of the graphic with overlays (roads/interstates and city landmarks) as riskier than the version without overlays, supporting H7.

4.6 Findings Related to Spatial Cognition

H8 predicted that numeracy and spatial cognition will be higher in the expert user group (EMs and meteorologists) than in the public sample. The numeracy values in the public sample were lower than in the expert group (Table 7). Spatial cognition was also lower in the public sample than in the expert group (Table 7).

Table 7. Descriptive statistics for numeracy [6-point scale] and spatial cognition [7-point scale] in both samples.

Variable	Sample	M	SD
Numeracy	Expert	5.05	0.64
	Public	4.08	1.02
Spatial Cognition	Expert	5.42	0.65
	Public	4.57	1.00

A Mann-Whitney U test was performed to evaluate whether numeracy differed by group (expert or public). The results indicated that the expert group was significantly more numerate than the general public group, $z = -4.23$, $p < 0.01$.

A Mann-Whitney U test was performed to evaluate whether spatial cognition differed by group (expert or public). The results indicated that the expert group had significantly higher spatial cognition than the general public group, $z = -3.89$, $p < 0.01$. Therefore, H8 was fully supported.

H9 predicts that numeracy and spatial cognition will predict the accuracy of graphic interpretation. In the expert user group, both qualitive and quantitative data was gathered for numeracy, spatial cognition, and accuracy. As outlined in the answer to H1, the expert user group was extremely accurate in reading the wind exceedance graphic. Additionally, as found in the answer to H8, experts were highly numerate and had high reported spatial cognition. One example of how experts would reflect on their abilities to interpret numbers and use maps accurately is the following:

> “Obviously, it's about a 10% chance of exceedance. That means 90% of the time, winds won't get this bad…so that means a lot to me as a meteorologist. To the public? Probably not as much. The conversion thing [knots to mph] is always interesting, but I understand the graphic… But seems to be fairly easy to understand, at least for me, from my perspective as a meteorologist” (Expert User 13).

In summary, in addition to experts’ high quantitative scores experts’ qualitative reflections point towards a positive relationship among numeracy, spatial cognition, and accuracy of graphic interpretation.

In the public sample, a multiple regression was conducted to test whether numeracy and spatial cognition predicted accuracy of interpreting the wind exceedance graphic. As a reminder, the average numeracy of the public was 4.08 ($SD = 1.02$) on a 6-point scale, average spatial cognition was 4.57 ($SD = 1.00$) on a 7-point scale, and the average accuracy score was 0.66 ($SD = 0.60$), where scores closer to zero are more accurate. Results from the multiple regression showed that numeracy did significantly predict accuracy ($\beta = -0.19$, $t(592) = -4.64$, $p < 0.01$).

However, spatial cognition did not predict accuracy ($\beta = -0.02$, $t(592) = 0.64$, $p = .53$), and the interaction between numeracy and spatial recognition was also nonsignificant ($\beta = -0.1$, $t(592) = -0.37$, $p = .71$). While numeracy was predictive of accuracy, neither spatial cognition nor the interaction was predictive; therefore, H9 is partially supported.

4.7 Expert Metacognitions of Vulnerability in Hurricane-prone Areas

The final research question (RQ4) wanted to know what metacognitions experts (EMs and meteorologists) had about vulnerability for their hurricane-prone areas. When experts prepare briefings, they think about their stakeholders in order to most effectively present their forecast information so the appropriate decisions can be made. By asking experts what a vulnerability in a hurricane is, I wanted to see how this group thought about what impacts would be present in their afflicted area and how other stakeholders may think about these impacts as well. Two major themes arose when expert users discussed vulnerability: physical vulnerability and social vulnerability. Physical vulnerabilities describe structures or systems with a increased potential for physical impact on the physical environment. Social vulnerabilities describe groups with an increased potential for social impacts from a disaster from individual and institutional systems. Each theme is discussed below with examples from the interviews.

4.7.1 Physical Vulnerability

When asked what vulnerabilities existed in their area, almost every expert participant (n = 17; 89.47%) responded initially by discussing the physical vulnerabilities, such as infrastructure and environmental features, that put their area at risk in a disaster. The most common physical vulnerabilities to hurricanes and hurricane-force winds mentioned were water drainage systems (e.g., Expert Users 1, 5, 6, 8, 10, 14, 17, 18), trees (e.g., Expert Users 1, 9, 14), power grid weaknesses (e.g., Expert Users 9, 10, 14, 15), and material aspects of mobile homes (e.g., 3, 6, 9, 12, 13, 15, 16). Experts reflected on how municipal systems and structures are critical to remain operational; if any of them go down during a hurricane, the aftermath can be deadly for people in the surrounding community (e.g., Expert Users 6, 9, 14).

The experts immediately thought about flooding connected to hurricanes. When natural and built water drainage systems become overwhelmed, flooding occurs and can occur well after the hurricane is over. Coastal storm surge and inland river flooding were brought up as the biggest physical threats from hurricanes (e.g., Expert Users 4, 5, 10, 12, 13, 14, 16, 17, 18). Flooding is a part of community memory, so it often gets the most attention during hurricanes. For example, Expert 3 described flooding from their personal experience:

> There's kind of this instinct reaction of 'hurricane means flood, and we need to start sandbagging'…when they don't really need to. I've seen that happen with Laura. We messaged Laura as a windstorm. And yet we had one official who was telling everybody to make sure they get sandbags. And so, so that mixed messaging—people remember things from the past event. And then they hold on to that and want to do that to take action, to prepare, when that may not be the action they need to take for this specific weather event (Expert User 3).

The infrastructure vulnerabilities in water and electricity are critical. Experts can tell how large the impact of the storm may be on an area based on the amount of time that passes from when power goes out and drainage systems stop working. These impacts are in the coastal zones as well as further inland and may last well into the recovery phase post-hurricane (Expert User 8).

The threat of fallen trees ties into the other physical vulnerabilities as well. Expert User 9 explained this relationship well:

> The number of trees is a big part of it. And a lot of them are knocked down…Trees blocking roads and taking out power is a big one. I mean, it's probably the biggest impact. Most folks are losing power due to downed trees and power lines. So as far as structures go, unless the tree falls on your home, most well-built single-family homes are going to withstand a tremendous amount of wind. You could start getting up into category three and higher, you know, you're going to have roofs blown off and things like that, but the home itself isn't going to necessarily get blown down (Expert User 9).

Expert 9, while initially discussing trees, brought up building codes like many of the other experts did (e.g., Expert Users 3, 13, 15). Experts talked about their communities' knowledge of building codes. For example, Expert User 15 was unsure of whether their community in northern Florida knew that there were different building codes in southern Florida counties.

Despite the inconsistencies in building codes within states, the single-family home is typically sturdier than the mobile or manufactured home. The vulnerability with mobile homes is a physical vulnerability because of the potential for structural damage. Hurricane-force winds will damage the mobile home parks physically. Expert User 13 stated, "There are a lot of mobile homes…very poorly built structures, we see it time and time again… you realize that it had no true foundation or anchoring. You know, you're very vulnerable, because the [building] codes aren't there." The inconsistencies in building codes for mobile home communities put residents at risk of their structures being blown away or collapsing, which they may not have the means to fix. This economic barrier to safety because of the physical risk is a large problem in the Gulf/Atlantic region. Expert User 3 explained:

> It's people that just don't have the means to evacuate or do anything to get them to a safer place. So if they live in, say, a mobile home, and they don't have a car, then they're kind of sitting like without a means to go anywhere to get to safety. So that's a big vulnerability for us in our area (Expert User 3).

Housing is more than physical vulnerabilities. Social vulnerabilities around housing include the ability to afford replacing physical housing structure, which are outlined in the following section.

4.7.2 Social Vulnerability

Despite an initial focus on physical vulnerabilities, participants in the interviews did bring up vulnerable populations in their forecasting areas and county boundaries (e.g., Expert Users 6, 7, 16, 17, 13). Three participants (15.79%) directly mentioned socio-economic and socio-demographic vulnerabilities as their first concern (Expert Users 2, 7, 19). For example, when asked what vulnerabilities were in their area, Expert User 19 responded:

> Vulnerabilities? I mean, I could talk about a lot of things. The main thing that I go towards…I always think about vulnerable populations. Louisiana is not a very rich state at all. And like, especially in like New Orleans, my main focus [is] public outreach…. a lot of these people don't get the education about these storms. They've lived through them, but no storm is the exact same.... Um, poor housing. Some people can't afford insurance, like flood insurance, house insurance, stuff like that (Expert User 19).

The ability to have insurance was mentioned specifically as a major vulnerability by five participants (26.32%). Experts emphasized that there was a gap in their communities surrounding knowing what type of insurance to have; insurance that covers flooding or that covers wind damage is different than general homeowner’s insurance (Expert Users 2, 10, 12, 18, 19).

Additionally, they reported that some residents could not afford multiple types of home insurance, even if they were aware of the types. Expert User 2 outlined these disparities:

> I think the biggest vulnerability we're going to have over the years is insurance... not being able to afford insurance, not knowing [what's covered] …We're rebuilding in the same coastal areas year after year after year, so money is really going to become a big problem in our coastal areas. So yeah, education, financial, and then just at what point are we going to stop building in some of these areas? (Expert User 2).

In addition to discussing lower socio-economic populations and the elderly (Expert User 6), other vulnerable populations that were discussed were those with language barriers (e.g., Expert Users 13, 19) and transient populations, such as snowbirds, tourists, or liveaboards[10] (e.g., Expert Users 2, 6, 7, 17). Language barriers, especially for groups that live in surge-prone areas, are necessary considerations; having evacuation orders or preparedness messages in Spanish or Vietnamese, for example, can make a large difference in recovery operations (Expert Users 6 and 13).

Transient populations, describing those who move frequently or are temporary to the area, are also a vulnerability in the Gulf/Atlantic region. As this part of the United States is warmer during winter, there is a significant population of "snowbirds" who move south until the weather gets more temperate up north (Expert Users 12, 16). However, as hurricane season stretches until late November when cooler temperatures begin in the northern United States, moving to the Gulf earlier in the winter months can place this group at risk because they may not know what to do during a hurricane. Tourists have a similar vulnerability, as they may choose to

[10] Snowbirds are people that live half the year in a warmer climate, typically moving south for the winter. Liveaboards are people who live full time on boats, such as a yacht, sailboat, or houseboat in a marina.

take vacations in the area without thinking about hurricane season and are likely to be without the necessary supplies to protect themselves in a rapidly intensifying storm (Expert Users 5, 12, 13). Lastly, liveaboard populations, or those that live on boats in ports year-round, were brought up as a vulnerable population that must be considered for sheltering and emergency assistance during hurricanes since their home's foundation is literally water (e.g., Expert Users 6, 7, 17).

RQ4 asked about what metacognitions expert users had of vulnerability for their hurricane-prone areas. Experts brought up physical and social vulnerabilities when describing their communities. Physical vulnerabilities were mentioned more than social vulnerabilities, including topics such as utility systems for water and power, fallen trees, and building codes. The social vulnerabilities discussed included socioeconomic and sociodemographic characteristics such as access to or knowledge of insurance, language barriers, and transient populations.

CHAPTER 5. DISCUSSION

There is no such thing as a *natural* disaster; disasters describe a disruption due to a hazard (environmental or anthropogenic) that has catastrophic impacts on the physical and social systems in a community. Hurricanes are extreme weather events that can become disasters for coastal and inland communities. Experts like EMs and meteorologists must effectively prepare for hurricanes and disseminate emergency messages to their areas to protect against the loss of life and property. Not only is there a need for effective communication, but experts have a desire to improve their messaging each storm. The importance of science communication for hurricane messaging has increased over the last 40 years as hurricanes have amplified in severity and coastal populations have grown (Eosco, 2008; National Oceanic and Atmospheric Administration, 2019; Ripberger et al., 2022). Within the last 10 years, there has been a push in the meteorological communication discipline to more rigorously study new hurricane graphics from a social-science approach (Morss et al., 2016; Reinhart et al., 2021). One of the key ways that hurricanes are communicated is through maps (Eosco, 2008; Ruginski et al., 2016). Therefore, studying maps as a communication tool for hurricanes can reveal how different experts and the public interpret hurricane messaging and reveal how experts think about their communities when using these graphics.

This dissertation is only the second time since the collective effort to test the storm surge graphic from 2011–2016 to study how hurricane map graphics are understood by experts and the public *before* the graphic is put into operations by the National Hurricane Center and National Weather Service forecasting offices (Morss et al., 2016). I designed this research to provide a model for how to plan a mixed-methods study to capture user design preferences, test future forecast graphics, and to add to the understanding of how the public and experts like emergency managers (EMs) and meteorologists interpret and perceive hazard map information. The findings are discussed below in order of importance to the field. Even if the wind exceedance graphic in this form does not get disseminated in future hurricane seasons, results from this study can be applied to different hazards. The mixed-methods approach afforded this study qualitative and quantitative insights into two important audiences for hurricane graphics—experts and the public. While there were some similarities between design preferences, how accurately each group interpreted the graphic and how the experts anticipated that the public would engage with the graphic differed. Sampling both audiences is important, demonstrating that design choices matter and influence how graphic information is perceived.

Previous social-science research in the severe weather domain has investigated how different components of map graphics, such as probability language and color choice, influence people's risk perceptions (Demuth et al., 2012; Morrow & Lazo, 2014; Morss et al., 2018; Ripberger et al., 2022). What has not been deliberately incorporated into forecast graphics are the principles of universal design (UD). UD describes the intentional incorporation of design features so that a message or product is easy to use, is appropriate for a wide range of consumers with different abilities, and communicates information effectively and equitably (General Services Administration, 2020; Hitt, 2018). This study connects UD to risk perception through a prototype hurricane graphic for wind exceedance. This research also looks at how the abilities of experts and the public, such as numeracy or spatial cognition, influence how accurately they interpret the graphic. As outlined by the RISP model, this study follows the influence of individual characteristics on decision-making for risk-related behaviors and extends what individual characteristics can be included (e.g., adding numeracy and spatial cognition) as well as suggests how UD can be used to support information processing to better support more audiences. Additionally, by asking how experts think about vulnerability when shown the graphic, this work extends UD thinking into the metacognition of disasters by revealing what experts consider to be more at-risk in their communities and what may need more support through messaging or resources in future hurricanes.

5.1 Accuracy, Numeracy, Spatial Cognition, and Confidence on Graphic Interpretation

Accuracy is the heart of research on risk graphics. The goal of hazard risk communication is to make graphics that are correctly understood so people can evaluate their risk and make decisions. The weather forecasting scholarship has focused around how users interpret information, centering around ability to decipher verbal or numerical information (Ripberger et al., 2022). However, many of the same influences, such as numeracy, confidence and spatial cognition, also influence how accurate someone is with interpreting risk graphics like maps (Hegarty et al., 2002; Hess et al., 2011; Roth, 2009). This study tested five hypotheses about the relationships between accuracy and individual influences (numeracy, confidence, and spatial cognition) to further the understanding of how people process information presented in a hurricane map graphic (H2, H3, H4, H8, H9).

Numeracy has been found to reflect a person's accuracy and confidence when they use numbers (Peters et al., 2019; Scott, 1999). As predicted, there was a relationship between numeracy and confidence in interpretation of the wind exceedance graphic (H4). Seen in the public sample, there was a moderate, positive relationship between numeracy and confidence. There was a positive, though nonsignificant, correlation in the expert sample between numeracy and confidence. The magnitude of these associations matters because it provides another way to descriptively focus on the relationships among the variables I studied. Whether the strength of association is looked at alone or in conjunction with statistical significance, the data suggests that there are connections between concepts and the value of the findings are not limited to just p-values. As the expert group is both small and more homogenous than the public group, the expert group's consistently high numeracy scores and confidence scores may explain the lack of statistical power to uphold the predicted relationship between variables, especially if the difference was small to detect. Additionally, while I did not find differences in the themes that arose from EMs and meteorologists in their interviews so I combined them into one expert sample, other studies have found differences between these two professionals. For example, the public and EMs have been found to be less numerate than meteorologists (Reinhart et al., 2021). Despite the statistical nonsignificance, finding a positive relationship in this study between numeracy and confidence affirms the relationship between these aspects found in the literature, which extends to accuracy.

Spatial cognition, defined as the ability to orient oneself within a map and comprehend its purpose, was incorporated into this study to measure a geographical equivalent to numeracy (Tomaszewski et al., 2020). As numeracy explains how people interpret numbers in risk messages, spatial cognition can explain more about how people interpret map graphics, specifically. Relationships between levels of spatial cognition and numeracy have been found, so it was anticipated that this relationship would be upheld within this study (Kane, 2014). The expert group was significantly more numerate than the general public group, again in line with previous studies on similar populations (Reinhart et al., 2021; Ripberger et al., 2022). The expert group also had significantly higher spatial cognition than the public group (H8). Adding spatial cognition measures into studies on forecast interpretation is new, only being incorporated into research recently (e.g., Davenport, 2022). This study provides an initial view into how measuring both spatial cognition and numeracy can elaborate into the characteristics of the populations of interest and how they interpret a hurricane graphic.

Bringing together the relationships between numeracy, spatial cognition, and accuracy, I hypothesized that numeracy and spatial cognition would predict the accuracy of interpreting the hurricane graphic (H9). The qualitative data from the expert interviews upheld this prediction; the expert user group was highly accurate, numerate, and had high reported spatial cognition. Results from the public sample also upheld the prediction; a multiple regression analysis revealed that numeracy and spatial cognition predicted accuracy. Interestingly, this analysis shows that spatial cognition did not significantly predict accuracy, though numeracy did. While there was not an interaction effect between spatial cognition and numeracy, the main effects show that when modeled together, accuracy can be predicted as a dependent variable. Knowing that one's levels of numeracy and spatial cognition can predict accuracy is valuable for future studies because it adds to our understanding of how information processing abilities can factor into how people interpret risk graphics. Additionally, knowing that numeracy is the stronger predictor between the two is valuable as well because it solidifies its importance as a variable, justifying how much of the literature on forecast risk message processing centers on numeracy.

The Risk Information Seeking and Processing (RISP) model is a systematic approach towards understanding how individuals perceive a hazard and ultimately perform information seeking behaviors (Griffin et al., 2013). Individual characteristics and hazard experience are major inputs in RISP to explain how people engage with hazard information. Demographic attributes have been found to influence information processing, such as education level (Yang et al., 2014). Individual characteristics that influence one's ability to process the hazard data information in the wind exceedance graphic (e.g., accuracy) in this study include numeracy and spatial cognition. Though educational level is already in RISP under demographics, having these more specific individual characteristics gave more context for how information was being processed. Additionally, I incorporated hurricane experience as a more context-specific measure for experience into this study because previous hazard experience has also been found to influence information processing within RISP (Yang et al., 2014). My findings reflected that individual characteristics and hazard experience influenced how people processed the hazard information in the wind exceedance graphic, supporting RISP's flow from these inputs. Also, because I had two groups, (a) a public that consumes the risk information and (b) experts who process and disseminate the risk information, I saw this flow with both groups. The experts in this study thought about their communities' characteristics and prior experiences and made assumptions of how they anticipated these groups would use the wind exceedance graphic. Thus, this study adds to RISP that risk information brokers (i.e., experts) follow the RISP flow when they think about their audience's processing of information when preparing hurricane graphics. As experts talked about wanting to make their messages easier to process for their communities, incorporating localization and UD into the graphic assists with this risk information flow.

5.2 Experience and Risk Perception

Experience with a hurricane has been studied from multiple standpoints around severe weather, from both personal and professional perspectives (e.g., Demuth, 2015; Dittmar & Duchin, 2016; Vasileiadou & Botzen, 2014). Experience has been linked limitedly to accuracy for interpreting maps (Hegarty, 2010; Smallman & Hegarty, 2007). Therefore, this study aimed to affirm the relationship between experience and accuracy, and to make connections to risk perception. As predicted in H1, the experts were extremely accurate in interpreting the wind exceedance graphic. Participants said that their extensive personal and job experience made it easier for them to accurately interpret the graphic, versus a novice user such as a member of the public (e.g., Expert Users 5, 11, 18). In the public survey, there was not a statistically significant relationship between accuracy and experience. Though non-significant, the direction of the correlation was positive, which adds to the literature on the relationship between experience and accuracy, though with limited statistical power (e.g., Smallman & Hegarty, 2007).

The experts in this study interact with similar graphics to the wind exceedance graphic as a routine aspect of their jobs. The experts had a lot of experience with hurricanes and with hurricane graphics, and so they could be more disposed to accurately interpret the graphic. The public has also seen similar graphics, but they only interact with the graphics when a hurricane is likely to impact their area. The public's lower exposure to and familiarity with these types of graphics could have factored into their low accuracy scores. Familiarity with a type of risk graphic could be an input into the RISP model, much like hazard experience. If people have less experience with a style of risk information (e.g., hurricane graphics), there could be a potential influence on their future information seeking or processing. Future RISP studies should consider adding experience with hazard graphics as an input into the model.

Hazard experience was measured with the Hurricane Experience Score (HES) in the public, and the score was low in this group. The HES measure used for experience in this study arguably did not capture personal experience as completely as it should have. The HES left out hurricane experiences during the forecasting period in the week lead-up to a potential landfall. Also, the scale excluded preparedness actions as well as media consumption before a hurricane, which is a major component of hurricane experience and could explain why the scores were so low in an area that routinely has hurricanes, especially in the last 30 years (National Oceanic and Atmospheric Administration, 2019). A more encompassing measure for hurricane experience from before the storm through recovery could elaborate more on how risk perception and experience are related for this specific hazard. I would suggest that future researchers create a more representative scale for hurricane experience, such as Demuth (2015) did for tornado experience.

Moving away from the ability to accurately use the risk graphic, I wanted to see how experience related with hurricane risk perception. Much like how other researchers have found that more experience with a hazard led to higher risk perceptions (e.g., Keller et al., 2006; Morss et al., 2016; Siegrist & Árvai, 2020), I was curious if there were differences between professionals and the public when it came to this relationship between experience and risk perception. RQ1a,b asked if experience affected risk perceptions in experts (RQ1a) and the public (RQ1b) because if the knowledge gained from direct or indirect participation or observation of a hazard influences how we engage with hazards (Vasileiadou & Botzen, 2014), there would likely be a difference between those that participate in hurricanes as part of their jobs and live in the area, versus those that just live in the area.

During the interviews conducted in this study, EMs and meteorologists (experts) frequently referred to their professional experience to explain how they process general hurricane risk. For example, they would list the number of major storms and years they spent in their job positions as evidence for why they may react more calmly to a forecast than someone else. Additionally, the experts would use their professional experience to explain why they may be more engaged with new forecast graphics or information than someone who does not routinely have this type of information come across their desks. While I asked about experience in general, the experts would discuss their professional experience and personal experience to provide context for how they thought about risk. Like the suggestion above for a more comprehensive hurricane experience scale to describe personal experience, it would be equally impactful if future researchers were to create a professional experience scale for experts like EMs and meteorologists to more exactly identify the role professional experience has in risk perception.

To supplement the qualitative reflections between experience and risk perception, I measured general risk perception for hurricanes (which captures severity, consequences, and affect) to see where experts fell. Experts' quantitative risk perception ratings for hurricanes were only slightly above the middle-point ($M = 3.67$, $SD = 0.51$). However, they would reflect on the variety of experiences they had from major landfalls to near-misses in their interviews, which may explain the average rates overall. Despite having less hurricane experience, surprisingly, the public was also similarly above the middle-point for risk perception of hurricanes ($M = 3.61$, $SD = 0.72$). In the public test, hurricane experience was only measured with the HES, as there was not an opportunity for an open-ended response. While the public's HES score was not that high on average, hurricane experience was a statistically significant predictor of risk perception. Despite the limitations with the HES, this study did find a relationship between experience and risk perception, supporting other previous studies about these variables (e.g., Siegrist & Árvai, 2020).

5.3 Perceptions and Impressions of Universal Design Elements in the Wind Exceedance Graphic

5.3.1 Color Preference and Perceived Risk in Experts and the Public

This study incorporated UD into the development of the wind exceedance graphic to formally test colorblind-friendly design choices for a hurricane hazard. The goal of integrating UD principles into communication is to craft messages for a wide range of users with different abilities so they can access the information effectively and equitably (Crameri et al., 2020;

General Services Administration, 2020; Hitt, 2018). Color is one of the most widely applied strategies used in risk communication graphics to influence an audience's attention and risk perceptions (Bostrom et al., 2008). Hurricane graphics commonly use rainbow color schemes, which are not only unreadable for those with color-vision deficiencies, but rainbow schemes distort data on graphics because certain colors can unfairly draw people's eyes to one part of the graphic rather than the whole (Crameri et al., 2020). Therefore, by testing audience preferences and perceptions of accessible color schemes, this study provides insight into how these color schemes work for hurricanes.

I wanted to learn what impressions experts and the public had of a reds-only, yellow-to-red, and a texture color scheme in the prototype graphic (RQ2a, RQ2b) because while much of the research on color focuses on how accurately graphics are used, it is important to understand why people may prefer one presentation over another. Preference is often overlooked when designing a new graphic because it is a subjective measure that does not necessarily complement objective measures, like accuracy. Design preference should be considered in the design process because if users do not like a design feature in a graphic, they will not use the graphic. As seen in the experts' reflections on the texture scheme, there were strong, negative reactions to the graphic and they asked me to immediately take it off the screen so they would no longer have to look at the texture (e.g., Expert Users 4, 13). If the texture scheme was not tested and instead just implemented into a forecast graphic, it is likely that many people would have a similar reaction and choose to simply go without that hazard information because they would prefer to not look at the graphic at all.

I also wanted to know how risky the color choices seemed to the user groups, extending the literature on the relationship between risk perception and color choice to universal design

(H5, H6). An interesting and unexpected finding from this research was that there was a mismatch between color scheme preference and how risky each group rated the color schemes. As a reminder, the public overwhelmingly preferred the yellow-to-red color scheme (67.15%). The reds-only color scheme was the next preferred (22.60%), followed by the texture scheme (10.26%). Echoed by the experts, 16 interview participants also preferred the yellow-to-red color scheme (84.21%), followed by three who preferred the reds-only color scheme (5.79%), and no one preferred the texture scheme. While both experts and the public preferred the yellow-to-red scheme, the experts found this scheme to be the riskiest while the public found the reds-only scheme to be the riskiest. In their interviews, experts said they preferred the risker (yellow-to-red) color scheme because the hazard seemed more urgent, and they expected the public to react similarly (e.g., Expert Users 8, 17). The survey did not have open-ended responses to find out why, but the public rated the red color scheme as significantly riskier than the yellow-to-red color scheme. While not having more explanation from the public is a limitation of this study, I expected that people would prefer a color scheme that was not as risky (e.g., prefer yellow-to-red instead of reds-only) because it presented as less threatening, perhaps stimulating a less scary affective response. More research is needed to learn about the public's preferences.

While each color scheme used in the prototypes has been shown to provoke a sense of risk in emergency map graphics (e.g., Bostrom et al., 2008; Cheong et al., 2016; Lindell et al., 2021), I wanted to see if there was a relationship between perceived riskiness with preference (H6). The relationship was clear in the expert sample; the majority of participants did prefer the yellow-to-red color scheme and perceived it as riskier than the other color schemes. In the public sample, however, the relationship between risk and preference was not as clear. The reds-only scheme was scored as the riskiest on average by every preference group. Reds-only was the riskiest for those who preferred the reds-only color scheme, those who preferred the yellow-to-red scheme, and by the texture preference group. When analyzed further, there was no statistically significant difference between color scheme preference groups on rating the riskiness of the color schemes, so this study cannot say at this point if there is a relationship between color preference and perceived riskiness. However, future research should continue to test this relationship to add to the understanding of where user design preference may factor into risk perception, and vice versa.

5.3.2 Overlay Preference and Perceived Risk in Experts and the Public

Just as hurricane graphics do not follow UD principles for color, most hurricane graphics do not include any overlays, with the exception of the storm surge graphic that includes highways (Morrow et al., 2015). Localization, or the adding of identifiable places or landmarks into maps, serves as a UD measure to increase the accessibility and ease of use for a graphic. Studies have shown that adding cartographic features onto risk map graphics, such as interstates/highways or cities, helps individuals visualize the spatial extent of a hazard (Henstra et al., 2019; Lindner et al., 2018; Retchless, 2014). The easier that someone can visualize a hazard, the better they can assess their risk. Just as with color, adding landmarks into map graphics has been found to influence how risky someone perceives a hazard to be, so studying the effect of landmarks with perceived risk adds to the understanding of how designing graphics to be more accessible for a wide audience can impact risk perception (Lindner et al., 2018).

Since overlays have not been comparatively tested in hurricane graphics, I wanted to measure user preference as well for two different overlays (interstates/highways and city landmarks) to find out how experts and the public perceived these design choices for the graphic (RQ3). Overall, I found that having overlays was preferred over not having any, though there were differences between experts and the public. The public preferred the city landmarks (42.47%), followed by the interstate/highway overlay (21.79%). Having no overlay was preferred by 19.71% of the sample, and 16.03% stated that they equally preferred the overlay options. The overlay preference was more spread out in the expert user group; 47.37% preferred the city landmarks, 42.11% preferred the interstate/highway overlay, and 10.53% preferred neither overlay. The experts who preferred the city overlays said this was because they thought the public could more easily and accurately know where they are on the map (e.g., Expert Users 2, 4, 5, 8, 13, 16). Additionally, by identifying major cities, the hazard map brought more attention to the population centers, which can be helpful for experts when they are briefing stakeholders where social impacts of hurricanes would likely be (e.g., Expert Users 4, 7, 18).

Experts who preferred the interstates/highways overlay gave similar reasons as those who preferred the cities. Experts said they thought the public could more easily orient themselves on the map with roads (e.g., Expert Users 5, 8, 9, 13). Additionally, they felt that the roads made it easier for them to communicate where hurricane wind threats would lie (e.g., Expert Users 2, 3, 4, 5, 6). One large difference between the reasons the experts preferred the interstates/highways overlay and the reasons experts preferred city landmarks was that having the interstate overlay clearly pointed out evacuation routes, which are important for residents and tourists (e.g., Expert Users 2, 7, 8, 9, 11, 17, 18, 19).

I also wanted to see if there was a difference in the perceived riskiness of the graphic with or without the map overlay (H7). Most current hurricane map graphics do not include overlays other than the hazard data, such as expected feet of storm surge values or the track probability (Demuth et al., 2012; Eosco, 2008; Morrow et al., 2015). Therefore, I wanted to find out if adding more detail to future map graphics takes away from the perceived risk from the data being presented. To start with the quantitative data, the public rated the versions of the graphic with overlays (roads/interstates and city landmarks) as riskier than the version without overlays (Table 6). There was no statistically significant difference between the interstates and city overlays. While a difference between types of overlays was not found, there was an impact on how risky the graphics appeared with the addition of the extra detail from the overlays. There is a limitation in this study because there were not open-ended responses to find out more about how risk was perceived with the overlays in the public, and future work should delve more into the qualitative insights of the public.

The interviews revealed more about the connections between risk and adding overlay detail onto the map graphic. The graphics with overlays prompted experts to think about the risk of their communities and what decisions they would make (e.g., evacuation) more than the versions without overlays. Experts discussed how incorporating more detail made the graphic more useful because it identified populations centers where people live, commute, and travel, where there is more at-risk (e.g., Expert Users 2, 5, 8). While useful for quick orientation on a map, finding that the overlays can also prime thinking about individual and community-level risk can be a valuable addition to a future graphic aimed at reaching a vast, impacted audience. While experts and the public may not agree on which overlay they prefer, having an overlay is desired by these groups and influences what they think about when using the graphic. Localization matters and should be studied in more depth in the future.

5.4 Vulnerability and Communicating Effectively to Diverse Audiences

Vulnerability describes the socio-cultural and environmental factors that can make populations more at-risk during disasters, which influences how and whether they engage with and process risk information (Griffin et al., 2013; Oliver-Smith & Hoffman, 2002; Wisner et al., 2004). While there have always been areas that are vulnerable to hurricanes, coastal communities are becoming more vulnerable and less resilient between each season as storms get stronger and more frequent with climate change (Millet et al., 2020; Peacock et al., 2011). Much of the scholastic understanding of vulnerability in hurricanes came as a response to Hurricane Katrina in 2005, centering around risk perception, decision-making, and communication between social groups in reaction to the weak response from emergency agencies (Marcus et al., 2006; Millet et al., 2020). While a solid foundation for establishing initial conceptualizations and measures for vulnerability, communities and disasters have changed in the last 20 years. The Intergovernmental Panel on Climate Change (IPCC) projects that extreme, highly active hurricane seasons will become more likely and impact vulnerable populations at an unforeseen degree (IPCC, 2023). I argue that having an updated account of how experts (EMs and meteorologists) think about vulnerability in their communities is long overdue.

Seen throughout the research questions of this study, the experts would often talk about what they thought the public would think when shown the wind exceedance graphic. Metacognition, meaning the feelings and beliefs about our thoughts and actions in any given scenario, has been applied to organizational leadership specifically to describe how individuals think during emergencies, as their job requires coordination and communication amongst diverse agencies and sectors (Huang & Yang, 2020; Kimrey, 2017; Marcus et al., 2006). I brought vulnerability into this study on graphics to see what experts' metacognitions of vulnerability for their communities are in their hurricane-prone areas when they are shown a risk map. An incredible finding was that not only do experts already think about vulnerability from a physical and social perspective when presented with a forecast map graphic, but they think about how vulnerable groups may apply the information from the graphic as well as what decisions they (as the expert) would need to make with this new information in their professional roles.

Past research on coastal communities in hurricanes have looked at vulnerability, though often focusing on either physical or social vulnerabilities (e.g., May, 2019; Niles & Contreras, 2019; Santos-Hernández et al., 2020). Physical and social vulnerabilities are both complex and vast in scope, so it makes sense that researchers would focus on either subject. In this study, however, I wanted to see what experts would bring up when asked just to identify and discuss vulnerability in their area without a prime of physical or social concepts. In the interviews, experts brought up both physical and social vulnerabilities. The first reaction of almost every expert (89.47%) was about physical vulnerabilities, such as infrastructure and environmental features. However, participants also discussed social vulnerabilities in their communities, such as transient populations or economic barriers to rebuild (e.g., Expert Users 6, 7, 16, 17, 13). In fact, three participants directly mentioned socio-economic and socio-demographic vulnerabilities as their first concern for the individuals and communities in their areas (Expert Users 2, 7, 19). The finding that experts thought about both types of vulnerabilities together is important, as it shows that this group naturally makes considerations for both aspects of vulnerability when they think about their areas during hurricanes.

The most common physical vulnerabilities to hurricanes and hurricane-force winds discussed were water drainage systems (e.g., Expert Users 1, 5, 6, 8, 10, 14, 17, 18), power grid weaknesses (e.g., Expert Users 9, 10, 14, 15), mobile homes (e.g., 3, 6, 9, 12, 13, 15, 16), and fallen trees (e.g., Expert Users 1, 9, 14). The impact and severity of utility outages and damage to mobile homes are also frequently discussed in other studies of physical vulnerabilities in hurricanes (e.g., Flood & Schechtman, 2014; Santos-Hernández et al., 2020). The prevalence of fallen trees in this study can be attributed to the graphic's focus on wind. The experts were concerned about the physical limitations some areas have to deal with flooding and other infrastructure weaknesses. Low-lying areas or outdated systems can have catastrophic impacts on the surrounding populations when they are damaged or inundated with water. The wrong decision to open shelters or conserve utilities can exacerbate the impact of physical vulnerabilities. Especially in the Gulf/Atlantic, the communities along the coast remember how devastating storms can be when emergency orders are not executed effectively (Expert Users 3, 8). The experts in this study recognized where these physical weaknesses were in their areas on the map graphic, especially in coastal regions and inland river waterways. It is useful to know that experts consider physical vulnerabilities both when thinking generally about hurricanes and when shown a hazard risk map for a wind threat. Additionally, experts think about how the physical impacts during hurricanes in these areas compound and limit the ability for people to thrive.

Social vulnerabilities arose from this discussion on one's ability to thrive, as there are municipal and individual economic considerations that have to be thought about during hurricanes as well. Socio-economic and socio-demographic vulnerabilities were brought up as examples for ways that certain populations have systemic constraints on their ability to bounce back or rebuild after a storm (Expert Users 2, 7, 19). For example, experts discussed how there were economic constraints for insurance in coastal communities. Not only are there differences between who can afford home insurance, but there are also major knowledge gaps between populations about the differences between types of home insurance (e.g., flood versus wind). Uninsured groups and mis-insured groups are extremely vulnerable, and the economic impacts on themselves and their communities affect these areas well through the recovery phase post-hurricane.

Beyond income, socio-cultural vulnerability was identified by the experts as well. There are transnational cultural considerations like what language is spoken and read at home (e.g., Expert Users 13, 19) that influence how people engage with emergency risk communication. If a major part of the community does not speak or read English, for example, map graphics may help some to understand the forecasted threat, but not all of the information may be accurately processed. Additionally, as more populations move into and out of coastal areas, the loss of local culture becomes a social vulnerability as well. Transient populations do not have the regional knowledge specific to an area that locals have, which makes them more vulnerable if they were to experience a hurricane since they are unfamiliar with what to do in this type of threat or how the area typically reacts to such an environmental disturbance (e.g., Expert Users 2, 6, 7, 17). Map graphics with localization features like overlays can help transient populations mitigate their risk, since there are identified features and landmarks that help them make more informed decisions during a hurricane in their new area.

Messaging was brought up by many of the experts as both a challenge and potential solution for protecting their vulnerable populations (e.g., Expert Users 3, 5, 6, 13). The experts in this study stressed how important it was for them as EMs and meteorologists to know how their communities understand emergency information throughout a hurricane. EMs and meteorologists communicate their forecasts and assessments to decision-makers, community partners, and the public, which have varying vulnerabilities and priorities. In hurricane messaging, it becomes necessary for EMs and meteorologists to have text products like evacuation orders or preparedness messages, as well as map graphics that can be understood and explained easily. The more that experts learn about the specific vulnerabilities within the communities in their forecasting area, the more effectively they can communicate. The experts in this study already were beginning to pair where the combination of physical and social vulnerabilities place groups more at risk in their communities, and they expressed a desire to keep learning about their people so that more people are kept safe in future hurricanes (e.g., Expert User 13).

Findings from this research show that experts such as EMs and meteorologists think about vulnerability in their areas when they think about hurricanes and when they are shown a hazard risk map. As revealed in the interviews, vulnerability is thought about both from a physical standpoint as well as a social standpoint. While this should not be surprising, it is encouraging to see in practice that these experts do have a robust understanding of vulnerability and have applied it to their unique communities. Additionally, these experts make audience considerations when they craft messages or prepare briefings to reach vulnerable audiences more effectively during emergencies like hurricanes. This study provides documented insight into how vulnerability is thought by experts when they are shown a hazard risk map, and future work should continue to delve into the metacognitions of experts for how the public sees their vulnerabilities during hurricanes.

5.5 Limitations

While this study adds to the literature on incorporating universal design principles into emergency graphics and the individual characteristics that influence how people understand map graphic information, the study has limitations. First, the 2022 hurricane season was unusually calm and may have influenced how both the experts and public were recalling their hurricane risk perceptions. Had it been a more normal, active season, there could have been higher risk perceptions since the area would have more recently gone through the preparedness to recovery activities for a hurricane. If people were only thinking about the current year, then there would not be much experience to draw from. Experts in the interviews would talk about more active years like 2020 rather than the current year because of the lack of any storms to discuss.

While the public sample was representative of residents of Florida and Louisiana, both states not having a landfalling hurricane (with the exception of Ian in the later season) was an atypical occurrence. There did not appear to be a history effect through the interviews with experts since they only mentioned Ian when they were listing their general experiences. Hurricane Ian may have influenced the survey respondents since it was a major hurricane that made landfall in Florida about a month before data collection. I could have added an item to the survey asking about their experience specifically in Ian to control for this influence in Florida.

Additionally, the calm 2022 hurricane season led to the cancellation of the observation method. If observations could have happened before the interviews and the survey, I could have gained additional insight about how hurricane graphics and vulnerability are discussed in live briefings. Additionally, the observation may have provided insight into the design of the most used or preferred hurricane graphics. This information would have been helpful because it would have given direct evidence of what graphics experts gravitate towards using, elaborating on preference. Also, since experts discussed how important it was for them to have the ability to communicate where impacts on the map were easily, attending briefings with graphics that did not have overlays would have elaborated more on how experts currently localize map graphics via their live explanations. Also, experts at the National Hurricane Center could have been recruited had we been introduced during an observation, which may have added to my understanding of how experts communicate to diverse audiences. Especially since the target audience of the NHC is at state level, rather than at the county level, having this potential pool of interviewees in the expert sample would have added to my understanding of how experts anticipated using and communicating the wind exceedance graphic to different audiences.

The sample size of the expert interviews is a limitation that should be expanded upon in future research. While I did reach saturation after 19 interviews, the total number of interviews was not representative of all experts in the emergency management and meteorological forecasting domains. While a qualitative study with a convenience sample cannot get the same representativeness as another method, future work should aim to capture the perspectives of more experts in other states to reflect the hurricane context more completely. Additionally, I interviewed meteorologists and EMs that were more senior in their positions. Only one expert was within the first couple years of their career, and most had been in their positions for 10 years or more. Especially as the workforce is changing and more EMs and meteorologists enter the career each year, capturing the perspectives of newer experts may reveal interesting information about how less experienced professionals interpret hazard maps and think about their communities.

Participants in both studies were recruited from only two states—Louisiana and Florida. While these states have some of the most active hurricane seasons in the United States, there are other coastal states along the Gulf Coast and Atlantic Coast that experience hurricanes as well. Expertise and perceptions from Texas or North Carolina, for example, potentially could have provided different findings in both the interviews and public survey. In order to get a more comprehensive take on new hurricane graphic products in the future, more coastal states should be sampled.

There are tensions between ways of knowing or ways of reporting evidence within the interdisciplinary research space. The value and challenge of this study was to have both qualitative and quantitative data. Studies with both types of data address research questions more wholly when incorporated together, though there are different approaches required when writing results (Hitchcock & Onwuegbuzie, 2020). However, there may have been a limitation in how the hypotheses and research questions of this dissertation were written in order to predict relationships with both types of evidence. For example, quantitative or statistical language was used sometimes to predict relationships that were tested with a qualitative method (interview) in order to meet the expectations of the meteorological discipline. Therefore, when I would report the findings, my hypotheses were either supported or unsupported because of how the prediction was written. In future research, I would use more open research questions for my expert sample, which mainly used qualitative data, and keep my predictive hypotheses for my public sample, which solely had quantitative measures.

Lastly, this study tested a new hurricane graphic for wind exceedance with experts and the public. At this point in time, I do not know if and how the public will use this graphic since it is new. However, at one point the time-of-arrival graphic and cone of uncertainty track graphic were new, and they turned out to be popular (Eosco, 2008). Thus, while it was valuable to gain their perspectives on color and overlays for a wind threat, the exceedance information itself may not have been the most applicable to them, which time will hopefully elaborate on. I learned from the informal interviews with the Hurricane Specialist Unit and the in-depth interviews with study participants that experts intend to use wind exceedance data in future forecasts, but were unsure how the public may use this information. Therefore, the graphic prototypes may have been processed more intentionally in the expert group because they anticipated needing to use it in the future, whereas I do not know how the public felt about wind exceedance.

CHAPTER 6. CONCLUSION

The consequences of a hurricane on a community are devastating, so experts like emergency managers (EMs) and meteorologists work to improve their hazard risk messaging each season. Map graphics are an increasingly popular means to communicate risk across an area. One reason for the popularity is that a map shows numerical, verbal, and pictorial information, which can communicate a wide amount of information in a single image. A good hurricane graphic is one that is accessible, accurately interpreted, and prompts users to think about what the data means. This study had the unique opportunity to follow a prototype wind exceedance graphic from development through testing to provide a model for how interdisciplinary researchers can design and assess new hurricane graphics. In a mixed methods approach, this study captured qualitatively and quantitatively user preferences and interpretations of a wind exceedance graphic.

This study was guided by parts of the Risk Information Seeking and Processing (RISP) model to explain how individual characteristics and hazard experience influence information processing with a risk graphic. One contribution of this study is that while many studies on RISP focus on how additional information is sought for a risk, this study uses RISP in the context of designing risk information. This dissertation contributes towards theory testing of RISP because the flow from individual characteristics to hazard experience to information processing was upheld, displaying a logical consistency between RISP and a context where this model is not used a lot, such as building risk graphics. Future studies on RISP and hurricane graphics should consider adapting the inputs of RISP to reflect the information design of risk graphics (e.g., previous experience with risk graphic, perceived hazard graphic characteristics, relevant channel beliefs where graphics are disseminated, and information seeking with additional graphics) and further test the explanatory power of the RISP model in the graphics context.

I incorporated universal design principles into the prototype graphics, focusing on how color scheme and cartographic overlays are perceived and influence the risk perceptions of experts and the public. Overall, the yellow-to-red color scheme was largely preferred by the expert group and the public sample. Though the expert group perceived this scheme as the riskiest, the public perceived this scheme as less risky than a reds-only color scheme. Experts preferred having contrasting colors and overlays in general on the graphic because it made communicating specifics about the wind threat easier, which is a major component of their job responsibilities during hurricanes. City landmarks were the most preferred overlay by the public, though experts almost equally preferred the interstate/highway overlay. Having cities or roads overlayed onto the graphic increased the perceived riskiness of the map graphic, potentially because it explicitly identified where major populations were.

Additionally, I was interested in seeing how individual characteristics and experience influenced how people processed the wind exceedance data from the map graphic, following the predictive flow of RISP. The experts had a large amount of experience working in hurricanes, high numeracy levels, and high spatial cognition abilities, which they felt made it easier for them to accurately interpret the new graphic than would the public. The public has less experience with hurricanes, lower numeracy levels, and lower spatial cognition scores than experts. Likewise, the public was less accurate in interpreting the graphic. Despite the differences between samples, there were relationships between these variables that helped to explain how people process and interpret forecast map graphics. Numeracy and spatial cognition predicted the accuracy of interpretation for the graphic. Though the effect was small, this study showed that the ability to use numbers influenced how accurately one interprets a probabilistic hurricane map graphic.

Lastly, I was interested in seeing how experts thought about vulnerability in their communities. During emergencies, there are different groups and infrastructures at-risk that can factor into how experts make decisions with forecast information. In the interviews, experts brought up physical and social vulnerabilities that were prevalent in their communities during hurricanes. Especially as vulnerability can be cumulative during disasters, seeing experts thought about municipal infrastructures, the environment, and socio-demographic attributes brought up how nuanced decisions need to be during extreme weather events. Overall, this study aimed to capture the impressions and use of a wind exceedance map graphic, which can hopefully be applied towards future meteorological graphic product development.

6.1 Future Research

Over the course of this dissertation, I learned a lot about how interdisciplinary teams can work together from the start of a project. The reason that the wind exceedance graphic was developed was not only because EMs and meteorologists had asked in previous seasons for this graphic, but also because CIRA and the NHC wanted to update the WTCM model and improve its windspeed forecasting. By bringing communication scholars and atmospheric scientists together at the start of this joint effort, universal design choices could be determined at the same time as the mathematical modelling was being written and the development of WTCM and design of the wind exceedance graphic could happen together.

A suggestion for future studies on the wind exceedance graphic is to conduct another interdisciplinary study on how experts and the public are actually using the graphic now that it is out of testing. This could be done through talk-aloud observations of EMs and meteorologists using the wind exceedance graphic in their briefings during an active hurricane season. This study focused on the designing and testing of the graphic before it was released for use during a hurricane season, so future studies can test the accuracy of interpretation and perceptions of the graphic once it is regularly used by experts. Social science has often been thought of as slower than operational meteorology, but this study showed that both types of research could be done concurrently and at a similar speed. Future graphic development should echo this interdisciplinary process, as there was a mutually beneficial outcome for both groups of scientists. The goal of both groups is the same: to improve hurricane graphics for their intended audiences.

Guided by the risk information seeking and processing (RISP) model, this study employed a more systematic approach to study how individual characteristics and hazard experience influenced risk perception and information processing of a new wind exceedance hurricane graphic. Numeracy and spatial cognition were measured in this research with other individual characteristics from RISP (E.g., demographics), as well as previous hurricane experience. My study was not interested in information seeking or information insufficiency because I gave the participants the information (i.e., the prototypes of the wind exceedance graphic) I wanted them to engage with. However, a future study that looks into informational subjective norms around hurricane forecasts would be beneficial to learn about how people engage and seek more information during a hurricane. Informational subjective norms, or an individual's belief that others think they should stay informed about a given risk, may be interesting in a forecasting context. The NWS and NHC are highly trusted by the public, even though forecasts for a hurricane can change a lot over the 5-8 day lead up to landfall (Rosen et al., 2021). If there is an expectation to know the forecast for a hurricane the entirety of the storm, and the risk graphics are designed to be understood more accurately following UD, then there could be a positive influence on information seeking to other NHC graphics. Studying hurricane risk information through the entirety of the RISP model could reveal more about where people to go for severe weather information and what graphics people gravitate towards using more than others.

Additionally, RISP covers both systematic processing and heuristic processing around information seeking (Griffin et al., 2013). This study covers systematic processing mainly, though there are other aspects of heuristic processing that pertain to severe weather, such as the need for people to make decisions under time pressure. While hurricanes have about 5-8 days of forecasting before they make landfall, people may need to process hurricane graphics faster in cases of rapid intensification. Also, people may not pay attention to information about an impending hurricane until it is closer to landfall. In future research, the wind exceedance graphic could be tested with prototypes made to support more heuristic processing as a hurricane nears landfall. For example, adding time-of-arrival information to the wind exceedance graphic could be tested to see how time and wind data on a map helps with decision-making.

As mentioned throughout the discussion, research (like forecasting) is imperfect. Where there are limitations in the measures or results, there are opportunities to improve in future work. Hurricane experience is one area especially that needs to be studied more. This study looked at personal experience at a more general level quantitatively and began to assess the nuances of professional experience from the qualitative interviews. Future studies should create and validate more specific scales for professional experience because of how influential professional experience can be on decision-making in the workplace (Vasileiadou & Botzen, 2014). Additionally, personal hurricane experience should have a more comprehensive scale because there are more aspects that describe what it means to go through a hurricane than what is currently in the HES.

There are other details to consider when adding overlays to hurricane graphics. As experts talked about in their interviews, an overlay of county boundaries could help them brief their partners more effectively. Additionally, experts suggested including natural landmarks. Lakes or rivers are routinely used by experts to orient themselves in their area, and they were missing on the prototype graphic. Also, future researchers should test whether adding the names or numbers of highways to the highway overlay makes it easier for different groups to localize the hazard risk information in the maps. There is a challenge to find the most effective balance of localization on a map overlay that adds detail, but does not make the graphic feel cluttered. Testing multiple overlays is important to see what option is the most preferred and what option increases the chance that people will accurately place themselves on the map. Localization matters, so it is worth the research time, effort, and resources to test different approaches in graphics.

The relationships between accuracy, spatial cognition, and numeracy should continue to be studied in the future. Especially since numeracy has been a major focus in much of the studies on probability communication, having a geographic equivalent should also be a focus for probabilistic forecast graphics. While spatial cognition and numeracy did not have an interaction effect in a regression model predicting accuracy in this study, I think this interaction would be present in a future study with a more known hazard. A study on storm surge, for example, may reveal this relationship between accuracy, spatial cognition, and numeracy because it is a hazard that has been introduced to the public longer than wind exceedance and is also depicted in graphics. Additionally, while this study measured accuracy at the ratio level, I did not calculate any metrics that would tell us if the public tended to overestimate or underestimate the wind threat. Knowing whether the public has a tendency towards over- or under-estimation could better inform EMs and meteorologists about the public's interpretations of hurricane graphics.

There is nuance in public perceptions of risk. For example, experts in this study wanted to use the yellow-to-red scheme because they thought its use would increase people's risk perceptions. However, the public preferred the yellow-to-red scheme, which they rated as less risky. I do not know from this study whether the public likes the yellow-to-red scheme because it appears less risky to them, or if there was another reason. If there was a relationship between preference and risk perception in the public, the less risky color scheme could have made them feel more hopeful that the consequences of the storm were not as severe. Future research should look into why the public preferred this color scheme over the reds-only color scheme. Additionally, future research should aim to capture why there was a mismatch between color preference and riskiness in experts and the public. A shorter survey with open-ended responses would likely capture this difference.

One major inconsistency I have noticed is in regard to the "reasonable worst-case scenario" language for the 1 in 10 (10%) likelihood title of the wind exceedance graphic. In the interviews, the EMs said they do not use the worst-case scenario phrasing, but the meteorologists said that the EMs do use it. The meteorologists said they use the worst-case scenario phrasing only because the EMs use it. In addition to this paradoxical loop, when asked in an open-ended question in the survey, the public was confused about the phrase's meaning. I read through the open-ended responses briefly and I see a potential affective influence that using "worst-case" language has on the reader, as well as increased difficulties in comprehension with using "worst-case scenario" as a qualifier or modifying phrase to accompany the 10% probability. This was a brief read through the open-ended data, and a more thorough analysis is needed. Future researchers should investigate what the public and experts think “wind exceedance” means and what the public thinks “reasonable worst-case scenario (10% chance)” means; this descriptive phrase has not been looked at in previous studies of probability language and it may be a point of confusion if it is used in future hurricane graphics.

Lastly, this study focused on expert metacognitions of vulnerability for the public. While these reflections elaborated on how experts saw vulnerabilities in their communities, I did not directly assess the vulnerability of the public in the survey method. I did collect information on participants’ length of time in state, general location, hurricane experience, and demographics (e.g., gender, age, education, income, race/ethnicity). This information could be combined into a measure of vulnerability. This new measure could be validated against the Social Vulnerability Index (SVI) in future research. The SVI comes from the Centers for Disease Control and Prevention (CDC) and the Agency for Toxic Substances and Disease Registry, which uses 16 U.S. census variables to help local officials identify communities that may need support before, during, or after disasters (Centers for Disease Control and Prevention, 2022). The SVI groups the 16 variables into four categories: socioeconomic status, household characteristics, racial and ethnic minority status, and housing type/transportation. Previous work has used the SVI to measure social vulnerabilities in communities and connect them to insurance coverage after hurricanes, but this work has not connected SVI scores to physical vulnerabilities (Peacock et al., 2011). In future papers, researchers should compare demographics from hurricane survey responses to Social Vulnerability Index census information with known physical vulnerabilities in the same areas that surveys are distributed.

6.2 Practical Recommendations

This dissertation provides initial design ideas for forecasters to apply to new hurricane graphics, as well as a sample study design for rigorously testing these graphics a season before they are put into operations. My broad recommendation around color scheme for hurricane graphic developers is to use the yellow-to-red scheme. This scheme was the most preferred by the EMs and meteorologists, as well as the public. In addition to being colorblind-friendly, the yellow-to-red scheme is the most similar option to color schemes being used by the NHC. Implementing the yellow-to-red color scheme would be a less dramatic change, which may be more readily accepted by the current NHC and NWS workforce.

As for localization overlays, I highly recommend incorporating these into all hurricane graphics. The extra detail provided did not clutter the map and was positively perceived by both experts and the public. Additionally, interstate networks and major city markers are available from census data and their code can be easily integrated into current graphic models. If the NHC or NWS needed to choose which overlay would be best, the public in this study preferred to have city landmarks on the graphic. The expert users were more split on overlay preference between cities and interstates, though they agreed that having overlays was a better practice overall.

If a map graphic is being developed solely for expert users and will not go to the public, using county boundaries should be explored as an overlay since it was highly requested and experts said it would make communicating easier during hurricane briefings. Experts told me that they would sometimes add county overlays themselves when preparing for a briefing, and having this overlay as a pre-programmed option would save them time. Another potential overlay to look into would be natural landmarks, such as lakes. In this study, the experts brought up that the prototypes were missing Lake Pontchartrain in Louisiana and Lake Okeechobee in

Florida, which they said they routinely used to help orient themselves. Especially as the WTCM model was being updated specifically to address inland windspeed reduction over bodies of water, incorporating large bodies of water onto the map graphic would help to explain the model updates as well as helping experts to localize the hazard.

This study incorporated universal design into the graphic by making the color schemes colorblind-friendly. This design choice is a solid effort towards making hurricane graphics meet 508 compliance requirements, which requires all United States federal electronic content to be accessible (General Services Administration, 2020). However, there are more efforts that need to be made to increase the accessibility of hurricane graphics because universal design principles do improve information processing and risk decision-making for the larger public because it makes graphics easier to understand. For example, future graphic producers need to make sure that all new prototype graphics work with assistive technology such as text-to-speech converters or screen magnifiers for those who are visually impaired. Alt text for the graphic can also be coded to match hurricane text products so the forecast messages are connected directly to the image.

While the experts I talked to had a good understanding of their communities due to their lengthy experience in their areas, there is additional training I recommend that would benefit EMs and meteorologists. For example, taking courses or seminars on risk communication, disaster anthropology, and cognitive psychology could boost their understanding of how people process information and why they take or avoid certain behaviors during hurricanes. Also, the added understanding of human behavior from these courses could be particularly important for new EMs and meteorologists, i.e., experts who do not have a wealth of professional experience to draw upon to help them understand their audiences. If adding training on these subjects can more quickly increase how EMs and meteorologists may understand their audiences, then they can more quickly start applying this information in their communication.

The benefit of this study was that I could study graphic perceptions and interpretations through a social-scientific lens alongside the data product development. Previous work on hurricane graphics have looked at how people understand information being presented, but have not asked why they prefer to use one graphic over another. Measuring preference should continue to be incorporated into future graphic development projects to learn why people gravitate towards certain data presentations over others. After development, the CIRA team could make changes and test the graphic before it was fully operational in an active hurricane season. I highly recommend involving interdisciplinary scientists from development through testing, as it can reveal interesting and important information at a far more helpful pace. Often, social scientists become a clean-up crew of sorts, being brought into projects after decisions have been made about how model information is going to be presented. In this study, designing the social science methodological instruments and graphics was being done while mathematical modelling was being developed for the wind exceedance data, and neither group felt restricted by the simultaneous work of the other discipline. There are boundless advantages to conducting social science and physical science in tandem, and all future interdisciplinary teams should plan to involve both types of scientists from the outset.

REFERENCES

Adams, T. M., Sanders, S. J., & Shivers-Williams, C. A. (2017). An Exploratory Examination Of the Impact of Uncertainty on Forecast Data Preference. *Global Aspect*, *7*(2), 11–15.

Aerts, J. C. J. H., Clarke, K. C., & Keuper, A. D. (2003). Testing Popular Visualization Techniques for Representing Model Uncertainty. *Cartography and Geographic Information Science*, *30*(3), 249–261. https://doi.org/10.1559/152304003100011180

Ahn, J., & Noh, G. (2020). Determinants of environmental risk information seeking: an emphasis on institutional trust and personal control. *Health, Risk & Socieity*, *22*(3–4), 214–230. https://doi.org/10.1080/13698575.2020.1813261

Alexander, D. E. (2004). Cognitive mapping as an emergency management training exercise. *Journal of Contingencies and Crisis Management*, *12*(4), 150–159. https://doi.org/10.1111/j.0966-0879.2004.00445.x

AOML Communications. (2022). *Researchers and Forecasters Team Up to Improve Forecasts in the New Hurricane and Ocean Testbed*. NOAA Atlantic Oceanic and Meteorological Laboratory. https://www.aoml.noaa.gov/news/new-hurricane-and-ocean-testbed/

Bathi, J. R., & Das, H. S. (2016). Vulnerability of Coastal Communities from Storm Surge and Flood Disasters. *International Journal of Environmental Research and Public Health*, *13*(2), 239. https://doi.org/10.3390/ijerph13020239

Bednarz, R. S., & Lee, J. (2011). The components of spatial thinking: empirical evidence. *Procedia - Social and Behavioral Sciences*, *21*, 103–107. https://doi.org/10.1016/j.sbspro.2011.07.048

Bernile, G., Nguyen, P., Bhagwat, V., & Kecskés, A. (2021). Are the risk attitudes of professional investors affected by personal catastrophic experiences? *Financial Management*, *50*, 455–486. https://doi.org/10.1111/fima.12328

Bisantz, A. M., Marsiglio, S. S., & Munch, J. (2005). Displaying Uncertainty: Investigating the Effects of Display Format and Specificity. *Human Factors*, *47*(4), 777–796. https://doi.org/10.1518/001872005775570916

Blevins, A. M. D. (2017). Participant Observation. In M. Allen (Ed.), *The SAGE Encyclopedia of Communication Research Methods* (pp. 1188–1190). SAGE Publications, Inc.

Bolton, M. J., & Ault, L. (2020). Weathering the Storms: Workplace Wellbeing, Mental Health, and the U.S. Meteorologist. *IJURCA: International Journal of Undergraduate Research & Creative Activities*, *12*(1). https://doi.org/http://doi.org/10.7710/2168-0620.0293

Bolton, M. J., Ault, L. K., Greenberg, D. M., & Baron-Cohen, S. (2018). Exploring the Human Side of Meteorology: A Brief Report on the Psychology of Meteorologists. *Journal of Operational Meteorology*, *06*(03), 23–32. https://doi.org/10.15191/nwajom.2018.0603

Bostrom, A., Anselin, L., & Farris, J. (2008). Visualizing seismic risk and uncertainty: A review of related research. *Annals of the New York Academy of Sciences*, *1128*, 29–40. https://doi.org/10.1196/annals.1399.005

Bourveau, T., & Law, K. K. F. (2021). Do Disruptive Life Events Affect How Analysts Assess Risk? Evidence from Deadly Hurricanes. *The Accounting Review*, *96*(3), 121–140. https://doi.org/10.2308/TAR-2016-0392

Broad, K., Leiserowitz, A., Weinkle, J., & Steketee, M. (2007). Misinterpretations of the "cone of uncertainty" in Florida during the 2004 hurricane season. *Bulletin of the American*

Meteorological Society, *88*(5), 651–667. https://doi.org/10.1175/BAMS-88-5-651
Bukvic, A., Smith, A., & Zhang, A. (2015). Evaluating drivers of coastal relocation in Hurricane Sandy affected communities. *International Journal of Disaster Risk Reduction*, *13*, 215–228. https://doi.org/10.1016/j.ijdrr.2015.06.008
Campbell, L., Gray, S., MacIntyre, T., & Stone, K. (2020). Literacy, numeracy and health and wellbeing across learning: Investigating student teachers' confidence. *International Journal of Educational Research*, *100*, 101532. https://doi.org/10.1016/j.ijer.2020.101532
Centers for Disease Control and Prevention. (2022). *CDC/ATSDR Social Vulnerability Index*. Agency for Toxic Substances and Disease Registry. https://www.atsdr.cdc.gov/placeandhealth/svi/index.html
Chandler, J., Rosenzweig, C., Moss, A. J., Robinson, J., & Litman, L. (2019). Online panels in social science research: Expanding sampling methods beyond Mechanical Turk. *Behavior Research Methods*, *51*, 2022–2038. https://doi.org/https://doi.org/10.3758/s13428-019-01273-7
Cheong, L., Bleisch, S., Kealy, A., Tolhurst, K., Wilkening, T., & Duckham, M. (2016). Evaluating the impact of visualization of wildfire hazard upon decision-making under uncertainty. *International Journal of Geographical Information Science*, *30*(7), 1377–1404. https://doi.org/10.1080/13658816.2015.1131829
Cheong, L., Kinkeldey, C., Burfurd, I., Bleisch, S., & Duckham, M. (2020). Evaluating the impact of visualization of risk upon emergency route-planning. *International Journal of Geographical Information Science*, *34*(5), 1022–1050. https://doi.org/10.1080/13658816.2019.1701677
Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (Second). Routledge. https://doi.org/https://doi.org/10.4324/9780203771587
Cramer, E. (2017). Interviews for Data Gathering. In M. Allen (Ed.), *The SAGE Encyclopedia of Communication Research Methods* (pp. 804–807). SAGE Publications, Inc. https://doi.org/10.4135/9781483381411
Crameri, F., Shephard, G. E., & Heron, P. J. (2020). The misuse of colour in science communication. *Nature Communications*, *11*(1), 1–10. https://doi.org/10.1038/s41467-020-19160-7
Davenport, C. E. (2022). *Characterizing the Growth of Spatial Thinking Abilities Across Meteorology Courses*. https://teaching.charlotte.edu/sites/teaching.charlotte.edu/files/media/sotl/2021 SOTL Davenport.pdf
De Elía, R., & Laprise, R. (2005). Diversity in interpretations of probability: Implications for weather forecasting. *Monthly Weather Review*, *133*(5), 1129–1143. https://doi.org/10.1175/MWR2913.1
Demuth, J. L. (2015). *Developing a valid scale of past tornado experiences* [Colorado State University]. https://mountainscholar.org/bitstream/handle/10217/170377/Demuth_colostate_0053A_13352.pdf?sequence=1
Demuth, J. L., Morss, R. E., Morrow, B. H., & Lazo, J. K. (2012). Creation and communication of hurricane risk information. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/BAMS-D-11-00150.1
Dillon, R. L., Tinsley, C. H., & Burns, W. J. (2014). Near-misses and future disaster preparedness. *Risk Analysis*, *34*(10). https://doi.org/10.1111/risa.12209

Dittmar, A., & Duchin, R. (2016). Looking in the Rearview Mirror: The Effect of Managers' Professional Experience on Corporate Financial Policy. *The Review of Financial Studies*, *29*(3), 565–602. https://doi.org/10.1093/rfs/hhv051

Dowling, P. (1997). *The Sociology of Mathematics Education: Mathematical Myths / Pedagogic Texts* (1st ed.). Routledge. https://doi.org/https://doi.org/10.4324/9780203486870

Dunwoody, S., & Griffin, R. J. (2015). Risk information seeking and processing model. In H. Cho, T. Reimer, & K. McComas (Eds.), *The SAGE handbook of risk communication* (pp. 102–116). Sage Publications.

Ehrlich, M., Harville, E., Xiong, X., Buekens, P., Pridjian, G., & Elkind-Hirsch, K. (2010). Loss of Resources and Hurricane Experience as Predictors of Postpartum Depression Among Women in Southern Louisiana. *Journal of Women's Health*, *19*(5). https://doi.org/10.1089/jwh.2009.1693

Eosco, G. (2008). A Study of Visual Communication: Cyclones, Cones, and Confusion. In *Cornell University: Vol. Masters Th* (Issue August).

Erdman, J. (2022). *It's The Atlantic Hurricane Season's Least Active Start In 30 Years*. The Weather Company'. https://www.wunderground.com/article/storms/hurricane/news/2022-08-22-atlantic-hurricane-season-quietest-start-in-30-years

Fagerlin, A., Zikmund-fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring Numeracy without a Math Test: Development of the Subjective Numeracy Scale. *Medical Decision Making*, 672–680. https://doi.org/10.1177/0272989X07304449

Faulkner, M. (2019). *The Impact of CEO Past Professional Experience and Social Capital on Corporate Policies and Firm Performance*. ProQuest Dissertations Publishing.

Ferme, E. (2018). A Comparison of STEM and non-STEM Teachers' Confidence in and Attitudes towards Numeracy. In J. Hunter, P. Perger, & L. Darragh (Eds.), *Proceedings of the 41st annual conference of the Mathematics Education Research Group of Australasia* (pp. 298–305). MERGA.

Ferreira, N., & König, A. C. (2014). Sample-Oriented Task-Driven Visualizations: Allowing Users to Make Better, More Confident Decisions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 571–580.

Finucane, M. L., Alhakami, A. L. I., Slovic, P., & Johnson, S. M. (2000). The Affect Heuristic in Judgments of Risks and Benefits. *Journal of Behavioral Decision Making*, *13*, 1–17. https://doi.org/https://doi.org/10.1002/(SICI)1099-0771(200001/03)13:1<1::AID-BDM333>3.0.CO;2-S

Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. (1978). How Safe Is Safe Enough? A Psychometric Study of Attitudes Towards Technological Risks and Benefits. *Policy Sciences*, *9*(2), 127–152. https://www.jstor.org/stable/4531720

Flood, S., & Schechtman, J. (2014). The rise of resilience: Evolution of a new concept in coastal planning in Ireland and the US. *Ocean and Coastal Management*, *102*(A), 19–31. https://doi.org/10.1016/j.ocecoaman.2014.08.015

Frels, R. K., & Onwuegbuzie, A. J. (2013). Administering Quantitative Instruments With Qualitative Interviews: A Mixed Research Approach. *Journal of Counseling & Development*, *91*(April). https://doi.org/10.1002/j.1556-6676.2013.00085.x

Gao, M., Liu, Y., & Shi, Y. (2020). Do people feel less at risk ? Evidence from disaster. *Journal of Financial Economics*, *138*(3), 866–888. https://doi.org/10.1016/j.jfineco.2020.06.010

General Services Administration. (2020). *Universal Design and Accessibility*. Sections508.Gov.

https://www.section508.gov/develop/universal-design/
Gersmehl, P. (2014). *Teaching Geography* (Third). Guilford Press.
Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, *10*(2), 486–489. https://doi.org/10.5812/ijem.3505
Goddard, S. (2017). *Disaster Preparedness Knowledge, Beliefs, Risk-Perceptions, and Mitigating Factors of Disaster Preparedness Behaviors of Undergraduate Students at a Large Midwest University* [A.T. Still Unversity]. http://search.proquest.com/docview/1936380967/
Golledge, R. G. (2002). The Nature of Geographic Knowledge. *Annals of the Association of American Geographers*, *92*(1), 1–14. https://doi.org/10.1111/1467-8306.00276
Griffin, R. J., Boerner, F., Ortiz, S., & Dunwoody, S. (2008). After the Flood: Anger, Attribution, and the Seeking of Information. *Science Communication*, *29*(3), 285–315. https://doi.org/10.1177/1075547007312309
Griffin, R. J., Dunwoody, S., & Yang, Z. J. (2013). Linking Risk Messages to Information Seeking and Processing. *Annals of the International Communication Association*, *36*(1), 323–362. https://doi.org/10.1080/23808985.2013.11679138
Grounds, M. A., Joslyn, S., & Otsuka, K. (2017). *Probabilistic Interval Forecasts : An Individual Differences Approach to Understanding Forecast Communication*. *2017*.
Guzman, G. G. (2017). *Household Income: 2016* (Issue September).
Hegarty, M. (2010). Components of Spatial Intelligence. In *The Psychology of Learning and Motivation* (Vol. 52, pp. 265–297). Elsevier Inc. https://doi.org/https://doi.org/10.1016/S0079-7421(10)52007-3
Hegarty, M., Burte, H., & Boone, A. P. (2018). Individual differences in large-scale spatial abilities and strategies. In D. R. Montello (Ed.), *Handbook of Behavioral and Cognitive Geography* (pp. 231–246). Edward Elgar Publishing. https://doi.org/https://doi.org/10.4337/9781784717544.00022
Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, *30*, 425–447.
Henstra, D., Minano, A., & Thistlethwaite, J. (2019). Communicating disaster risk? An evaluation of the availability and quality of flood maps. *Natural Hazards and Earth System Sciences*, *19*(1), 313–323. https://doi.org/10.5194/nhess-19-313-2019
Hess, R., Visschers, V. H. M., Siegrist, M., & Keller, C. (2011). How do people perceive graphical risk communication? The role of subjective numeracy. *Journal of Risk Research*, *14*(1), 47–61. https://doi.org/10.1080/13669877.2010.488745
Hitchcock, J. H., & Onwuegbuzie, A. J. (2020). Developing Mixed Methods Crossover Analysis Approaches. *Journal of Mixed Methods Research*, *14*(1), 63–83. https://doi.org/10.1177/1558689819841782
Hitt, A. (2018). Foregrounding Accessibility Through (Inclusive) Universal Design in Professional Communication Curricula. *Business and Professional Communication Quarterly*, *81*(1), 52–65. https://doi.org/10.1177/2329490617739884
Houston, D., Cheung, W., Basolo, V., Feldman, D., Matthew, R., Sanders, B. F., Karlin, B., Schubert, J. E., Goodrich, K. A., & Contreras, S. (2019). The Influence of Hazard Maps and Trust of Flood Controls on Coastal Flood Spatial Awareness and Risk Perception. *Environment and Behavior*, *51*(4), 347–375. https://doi.org/10.1177/0013916517748711

Huang, Y., & Yang, C. (2020). A Metacognitive Approach to Reconsidering Risk Perceptions and Uncertainty: Understand Information Seeking During COVID-19. *Science Communication*, *42*(5), 616–642. https://doi.org/10.1177/1075547020959818

IPCC. (2023). Summary for Policymakers. In H. Lee & J. Romero (Eds.), *Climate Change 2023: Synthesis Report. A Report of the Intergovernmental Panel on Climate Change* (pp. 1–36). Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

Johnson, S. G. B., Merchant, T., Keil, F. C., & Johnson, S. (2019). Belief digitization: Do we treat uncertainty as probabilities or as bits? *Journal of Experimental Psychology: General*, *284*, 1–37.

Joslyn, S., Berg, L. N., & Nichols, R. M. (2009). Probability of Precipitation: Assessment and Enhancement of End-User Understanding. *Bulletin of the American Meteorological Society*, *February*, 185–194. https://doi.org/10.1175/2008BAMS2509.1

Joslyn, S., & Savelli, S. (2010). Communicating forecast uncertainty: Public perception of weather forecast uncertainty. *Meteorological Applications*. https://doi.org/10.1002/met.190

Kane, P. J. . (2014). *An investigation into estimation and spatial sense as aspects of workplace numeracy: A case study of recycling and refuse operators within a situated learning model*. Auckland University of Technology.

Keller, C., Siegrist, M., & Gutscher, H. (2006). The Role of the Affect and Availability Heuristics in Risk Communication. *Risk Analysis*, *26*(3), 631–639. https://doi.org/10.1111/j.1539-6924.2006.00773.x

Kimrey, L. C. M. (2017). Metacognitive Decision Making in Oil Spill Repsonse Behavioral Bias in Relation to Percieved Risk. *International Oil Spill Conference Proceedings*, 1453–1470. https://doi.org/https://doi.org/10.7901/2169-3358-2017.1.1453

Kinkeldey, C., Maceachren, A. M., Riveiro, M., Schiewe, J., Kinkeldey, C., Maceachren, A. M., Riveiro, M., & Schiewe, J. (2017). Evaluating the effect of visually represented geodata uncertainty on decision-making: systematic review , lessons learned , and recommendations. *Cartography and Geographic Information Science*, *44*(1), 1–21. https://doi.org/10.1080/15230406.2015.1089792

Lee, J., & Bednarz, R. (2012). Components of Spatial Thinking: Evidence from a Spatial Thinking Ability Test. *Journal of Geography*, *111*(1), 15–26. https://doi.org/10.1080/00221341.2011.583262

Lee, S., & Lee, D. K. (2018). What is the proper way to apply the multiple comparison test? *Korean Journal of Anesthesiology*, *71*(5), 353–360. https://doi.org/https://doi.org/10.4097/kja.d.18.00242

Lenhardt, E. D., Cross, R. N., Krocak, M. J., Ripberger, J. T., Ernst, S. R., Silva, C. L., & Jenkins-Smith, H. C. (2020). How likely is that chance of thunderstorms? A study of how national weather service forecast offices use words of estimative probability and what they mean to the public. *Journal of Operational Meteorology*, *8*(5). https://doi.org/10.15191/nwajom.2020.0805

Light, A., & Bartlein, P. J. (2004). The end of the rainbow? color schemes for improved data graphics. *Eos*, *85*(40). https://doi.org/10.1029/2004EO400002

Lindell, M. K., Bostrom, A., Goltz, J. D., & Prater, C. S. (2021). Evaluating hazard awareness brochures: Assessing the textual, graphical, and numerical features of tsunami evacuation products. *International Journal of Disaster Risk Reduction*, *61*(March). https://doi.org/10.1016/j.ijdrr.2021.102361

Lindell, M. K., & Hwang, S. N. (2008). Households ' Perceived Personal Risk and Responses in a Multihazard Environment. *Risk Analysis*, *28*(2). https://doi.org/10.1111/j.1539-6924.2008.01032.x

Lindell, M. K., & Perry, R. W. (2012). The Protective Action Decision Model: Theoretical Modifications and Additional Evidence. *Risk Analysis*, *32*(4), 616–632. https://doi.org/10.1111/j.1539-6924.2011.01647.x

Lindner, B. L., Johnson, J., Alsheimer, F., Duke, S., Miller, G. D., & Evsich, R. (2018). Increasing Risk Perception and Understanding of Hurricane Storm Tides Using an Interactive, Web-Based Visualization Approach. *Journal of Coastal Research*, *34*(6), 1484–1498. https://doi.org/10.2112/JCOASTRES-D-17-00222.1

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General Performance on a Numeracy Scale among Highly Educated Samples. *Medical Decision Making*, *21*, 37–44.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as Feelings. *Psychological Bulletin*, *127*(2), 267–286. https://doi.org/10.1037/0033-2909.127.2.267

Maceachren, A. M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., Hetzler, E., Maceachren, A. M., Robinson, A., Hopper, S., & Gardner, S. (2005). Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartography and Geographic Information Science*, *32*(3), 139–160. https://doi.org/10.1559/1523040054738936

Maceachren, A. M., Roth, R. E., Brien, J. O., Li, B., Swingley, D., & Gahegan, M. (2012). Visual Semiotics & Uncertainty Visualization: An Empirical Study. *EEE Transactions on Visualization and Computer Graphics*, *18*(12), 2496–2505.

Marcus, L. J., Dorn, B. C., & Henderson, J. M. (2006). Meta-Leadership and National Emergency Preparedness: A Model to Build Government Connectivity. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, *4*(2), 128–134. https://doi.org/https://doi.org/10.1089/bsp.2006.4.128

May, C. K. (2019). Resilience, vulnerability, & transformation: Exploring community adaptability in coastal North Carolina. *Ocean and Coastal Management*, *169*, 86–95. https://doi.org/10.1016/j.ocecoaman.2018.12.007

Millet, B., Carter, A. P., Broad, K., Cairo, A., Evans, S. D., & Majumdar, S. J. (2020). Hurricane risk communication: Visualization and behavioral science concepts. *Weather, Climate, and Society*, *12*(2), 193–211. https://doi.org/10.1175/WCAS-D-19-0011.1

Monmonier, M. (2008). Web Cartography and the Dissemination of Cartographic Information about Coastal Inundation and Sea Level Rise. In M. P. Peterson (Ed.), *International Perspectives on Maps and the Internet* (pp. 49–71). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-72029-4_4

Morrow, Betty H., Lazo, J. K., Rhome, J., & Feyen, J. (2015). Improving storm surge risk communication: Stakeholder perspectives. *Bulletin of the American Meteorological Society*, *96*(1), 35–48. https://doi.org/10.1175/BAMS-D-13-00197.1

Morrow, Betty Hearn, & Lazo, J. K. (2014). Coastal emergency managers' preferences for storm surge forecast communication. *Journal of Emergency Management*, *12*(2), 153–160. https://doi.org/10.5055/jem.2014.0169

Morss, R. E., Cuite, C. L., Demuth, J. L., Hallman, W. K., & Shwom, R. L. (2018). Is storm surge scary? The influence of hazard, impact, and fear-based messages and individual differences on responses to hurricane risks in the USA. *International Journal of Disaster Risk Reduction*, *30*(January), 44–58. https://doi.org/10.1016/j.ijdrr.2018.01.023

Morss, R. E., Demuth, J. L., & Lazo, J. K. (2008). Communicating Uncertainty in Weather Forecasts : A Survey of the U . S . Public. *Weather and Forecasting*, *23*, 974–991. https://doi.org/10.1175/2008WAF2007088.1

Morss, R. E., Demuth, J. L., Lazo, J. K., Dickinson, K., Lazrus, H., & Morrow, B. H. (2016). Understanding public hurricane evacuation decisions and responses to forecast and warning messages. *Weather and Forecasting*, *31*(2), 395–417. https://doi.org/10.1175/WAF-D-15-0066.1

Morss, R. E., Lazo, J. K., & Demuth, J. L. (2010). Examining the use of weather forecasts in decision scenarios: Results from a us survey with implications for uncertainty communication. *Meteorological Applications*, *17*(2), 149–162. https://doi.org/10.1002/met.196

Nadav-greenberg, L., & Joslyn, S. L. (2008). The Effect of Uncertainty Visualizations on Decision Making in Weather Forecasting. *Journal of Cognitive Engineering and Decision Making*, *2*(1), 24–47. https://doi.org/10.1518/155534308X284354.

National Academy of Sciences. (2006). Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts. In *National Research Council Committee on Estimating and Communicating Uncertainty in Weather and Climate Forecasts* (Vol. 129). https://doi.org/10.1017/CBO9781107415324.004

National Environmental Satellite and Data Information Services. (2022). *Hurricane Ian's Path of Destruction*. NOAA. https://www.nesdis.noaa.gov/news/hurricane-ians-path-of-destruction

National Hurricane Center. (n.d.). *Glossary of NHC Terms*. National Oceanic and Atmospheric Administration. Retrieved March 12, 2023, from https://www.nhc.noaa.gov/aboutgloss.shtml

National Hurricane Center, & Central Pacific Hurricane Center. (2022). *Definition of the NHC Track Forecast Cone*. National Oceanic and Atmospheric Administration. https://www.nhc.noaa.gov/aboutcone.shtml

National Oceanic and Atmospheric Administration. (2019). *Atlantic high-activity eras: What does it mean for hurricane season?* US Department of Commerce. https://www.noaa.gov/stories/atlantic-high-activity-eras-what-does-it-mean-for-hurricane-season

National Oceanic and Atmospheric Administration. (2020). *2020 Atlantic Hurricane Season takes infamous top spot for busiest on record*. US Department of Commerce. https://www.noaa.gov/news/2020-atlantic-hurricane-season-takes-infamous-top-spot-for-busiest-on-record

National Weather Service. (n.d.). *Impact-Based Decision Support Services (IDSS)*. National Oceanic and Atmospheric Administration. Retrieved April 18, 2023, from https://www.weather.gov/about/idss

Niles, S., & Contreras, S. (2019). Social Vulnerability and the Role of Puerto Rico's Healthcare Workers after Hurricane Maria. *Natural Hazards Center Quick Response Grant Report Series*, *288*(July), 1–9.

Norris, F. H., Perilla, J. L., Riad, J. K., Kaniasty, K., & Lavizzo, E. A. (1999). Stability and Change in Stress, Resources, and Psychological Distress. *Anxiety, Stress and Coping*, *12*, 363–396. https://doi.org/https://doi.org/10.1080/10615809908249317

O'Neill, A. (2022). *Age distribution in the United States from 2010 to 2020*. Statista. https://www.statista.com/statistics/270000/age-distribution-in-the-united-states/

Oliver-Smith, A., & Hoffman, S. (2002). Introduction. In J. Currey (Ed.), *Catastrophe & Culture: The Anthropology of Disaster* (pp. 1–22). Oxford.

Peacock, W. G., Brody, S. D., & Highfield, W. (2005). Hurricane risk perceptions among Florida's single family homeowners. *Landscape and Urban Planning*, *73*(2–3), 120–135. https://doi.org/10.1016/j.landurbplan.2004.11.004

Peacock, W. G., Grover, H., Mayunga, J., Zandt, S. Van, Brody, S. D., & Kim, H. J. (2011). *The Status and Trends of Population Social Vulnerabilities along the Texas Coast with special attention to the Coastal Management Zone and Hurricane Ike: The Coastal Planning Atlas and Social Vulnerability*. http://hrrc.arch.tamu.edu/media/cms_page_media/558/11-02R.pdf

Peer, E., Rothschild, D., Evernden, Z., Gordon, A., & Damer, E. (2021). MTurk , Prolific or panels? Choosing the right audience for online research MTurk , Prolific or panels? Choosing the right audience for online research. *SSRN Electronic Journal*. https://doi.org/http://dx.doi.org/10.2139/ssrn.3765448

Peters, E., Kate, M., Knoll, M. A. Z., Ardoin, S. P., Shoots-reinhard, B., & Simon, A. (2019). Despite high objective numeracy, lower numeric confidence relates to worse financial and medical outcomes. *PNAS*, *116*(39), 19386–19391. https://doi.org/10.1073/pnas.1903126116

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. ., Mazzocco, K., & Dickert, S. (2006). Numeracy and Decision Making. *Pyschological Science*, *17*(5), 407–413. https://doi.org/https://doi.org/10.1111/j.1467-9280.2006.01720.x

Reinhart, A., Washington, T., & Schauer, J. (2021). Tropical Roadmap Part 3 - Roadmap Brainstorm. *2021 NOAA Hurricane Conference*.

Renn, O. (2011). The social amplification/attenuation of risk framework: Application to climate change. *Wiley Interdisciplinary Reviews: Climate Change*, *2*(2), 154–169. https://doi.org/10.1002/wcc.99

Retchless, D. P. (2014). Sea level rise maps: How individual differences complicate the cartographic communication of an uncertain climate change hazard. *Cartographic Perspectives*, *77*(77), 5–20. https://doi.org/10.14714/CP77.1235

Ripberger, J. T., Bell, A., Fox, A., Forney, A., Livingston, W., Gaddie, C., Silva, C., & Jenkins-Smith, H. (2022). Communicating Probability Information in Weather Forecasts: Findings and Recommendations from a Living Systematic Review of the Research Literature. *Weather, Climate, and Society*. https://doi.org/10.1175/wcas-d-21-0034.1

Riveiro, M. (2016). Visually supported reasoning under uncertain conditions: Effects of domain expertise on air traffic risk assessment. *Spatial Cognition & Computation*, *16*(2), 133–153. https://doi.org/10.1080/13875868.2015.1137576

Rosen, Z., Krocak, M. J., Ripberger, J. T., Cross, R., Lenhardt, E., Silva, C. L., & Jenkins-Smith, H. C. (2021). Communicating Probability Information in Hurricane Forecasts: Assessing Statements that Forecasters Use on Social Media and Implications for Public Assessments of Reliability. *Journal of Operational Meteorology*, *9*(7), 89–101. https://doi.org/10.15191/nwajom.2021.0907

Roszkowski, M. J., Davey, G., & Grable, J. E. (2005). Insights from Psychology and Psychometrics on Measuring Risk Tolerance. *Journal of Financial Planning*, *18*(4), 66–77.

Roth, R. E. (2009). The impact of user expertise on geographic risk assessment under uncertain conditions. *Cartography and Geographic Information Science*, *36*(1), 29–43. https://doi.org/10.1559/152304009787340160

Rubin, C. B., & Cutter, S. L. (Eds.). (2020). *U.S. Emergency Management in the 21st Centure: From Disaster to Catastrophe*. Routledge.

Ruginski, I. T., Boone, A. P., Padilla, L. M., Liu, L., Heydari, N., Kramer, H. S., Hegarty, M., Thompson, W. B., House, D. H., & Creem-Regehr, S. H. (2016). Non-expert interpretations of hurricane forecast uncertainty visualizations. *Spatial Cognition and Computation*, *16*(2), 154–172. https://doi.org/10.1080/13875868.2015.1137577

Ruisch, B. C., & Stern, C. (2021). The Confident Conservative: Ideological Differences in Judgment and Decision-Making Confidence. *Journal of Experimental Psychology: General*, *150*(3), 527–544. https://doi.org/10.1037/xge0000898

Salovich, N. A., Donovan, A. M., Hinze, S. R., & Rapp, D. N. (2021). Can confidence help account for and redress the effects of reading inaccurate information? *Memory and Cognition*, *49*(2), 293–310. https://doi.org/10.3758/s13421-020-01096-4

Santos-Hernández, J., Méndez-Heavilin, A., Álvarez-Rosario, G. (2020). Hurricane María in Puerto Rico: Pre-Existing Vulnerability and Catastrophic Outcomes. In S. Cutter & C. Rubin (Eds.), *U.S. Emergency Management in the 21st Century: From Disaster To Catastrophe*. Routledge.

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The Role of Numeracy in Understanding the Benefit of Screening Mammography. *Annals of Internal Medicine*, *127*(11), 1023–1028.

Scott, D. (1999). Essential Ingredients for Numeracy. *Australian Primary Mathematics Classroom*, *4*(1), 4–8. https://doi.org/https://search.informit.org/doi/10.3316/informit.461588318002419

Shukla, J. (2016). Extreme Weather: Mental Health Challenges and Community Response Strategies. In S. L. Steinberg & W. A. Sprigg (Eds.), *Extreme Weather, Health, and Communities* (pp. 245–270). Springer International Publishing. https://doi.org/10.1007/978-3-319-30626-1_11

Siegrist, M., & Árvai, J. (2020). Risk Perception: Reflections on 40 Years of Research. *Risk Analysis*, *40*, 2191–2206. https://doi.org/10.1111/risa.13599

Slovic, P. (1987). Perception of risk. *Science*, *236*(4799), 280–285.

Slovic, P., & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science*, *15*(6), 322–325. https://doi.org/10.1111/j.1467-8721.2006.00461.x

Slovic, P., Peters, E., Finucane, M. L., & Macgregor, D. G. (2005). Affect, Risk, and Decision Making. *Health Psychology*, *24*(4), S35–S40. https://doi.org/10.1037/0278-6133.24.4.S35

Smallman, H. S., & Hegarty, M. (2007). Expertise, spatial ability and intuition in the use of complex visual displays. *Proceedings of the Human Factors and Ergonomics Society*, *1*, 200–204. https://doi.org/10.1177/154193120705100412

Smith, S. M., Roster, C. A., Golden, L. L., & Albaum, G. S. (2016). A multi-group analysis of online survey respondent data quality: Comparing a regular USA consumer panel to MTurk samples. *Journal of Business Research*, *69*(8), 3139–3148. https://doi.org/10.1016/j.jbusres.2015.12.002

Strauss, A. L., & Corbin, J. M. (1998). *Basics of qualitative research: techniques and procedures for developing grounded theory* (2nd ed.). Sage Publications.

Taekeuchi, Y. (1992). Sense of direction and its relationship with geographical orientation, personality traits and mental ability. *The Japanese Journal of Educational Psychology*, *40*, 47–53. https://doi.org/10.5926/jjep1953.40.1_47

Tanner, A., & Doberstein, B. (2015). Emergency preparedness amongst university students. *International Journal of Disaster Risk Reduction*, *13*, 409–413. https://doi.org/10.1016/j.ijdrr.2015.08.007

Tarp, K. (n.d.). *NOAA Hazardous Weather Testbed*. NOAA National Severe Storms Laboratory. Retrieved March 10, 2022, from https://www.nssl.noaa.gov/news/factsheets/hwt.pdf

Thompson, M. A., Lindsay, J. M., & Gaillard, J. C. (2015). The influence of probabilistic volcanic hazard map properties on hazard communication. *Journal of Applied Volcanology*, *4*(1). https://doi.org/10.1186/s13617-015-0023-0

Thorndyke, P. W., & Goldin, S. E. (1981). *Ability Differences and Cognitive Mapping Skill*.

Tomaszewski, B., Walker, A., Gawlik, E., Lane, C., Williams, S., Orieta, D., McDaniel, C., Plummer, M., Nair, A., Jose, N. S., Terrell, N., Pecsok, K., Thomley, E., Mahoney, E., Haberlack, E., & Schwartz, D. (2020). Supporting disaster resilience spatial thinking with serious GeoGames: Project Lily Pad. *ISPRS International Journal of Geo-Information*, *9*(6). https://doi.org/10.3390/ijgi9060405

Trumbo, C., Meyer, M. A., Marlatt, H., Peek, L., & Morrissey, B. (2014). An Assessment of Change in Risk Perception and Optimistic Bias for Hurricanes Among Gulf Coast Residents. *Risk Analysis*, *34*(6), 1013–1024. https://doi.org/10.1111/risa.12149

Turgut, M. (2015). Development of the spatial ability self-report scale (SASRS): reliability and validity studies. *Quality and Quantity*, *49*(5), 1997–2014. https://doi.org/10.1007/s11135-014-0086-8

Tversky, A., & Kahneman, D. (1993). Probabilistic Reasoning. In *Readings in philosophy and cognitive science* (pp. 1124–1131).

United States Census Bureau. (2021a). *QuickFacts: Florida*. https://www.census.gov/quickfacts/FL

United States Census Bureau. (2021b). *QuickFacts: Louisiana*. https://www.census.gov/quickfacts/LA

Vasileiadou, E., & Botzen, W. J. W. (2014). Communicating adaptation with emotions: The role of intense experiences in raising concern about extreme weather. *Ecology and Society*, *19*(2), 1–36.

Weinstein, N. (1989). Effects of Personal Experience on Self-Protective Behavior. *Psychological Bulletin*, *105*(1), 31–50.

Weisberg, S. M., Schinazi, V. R., Newcombe, N. S., Shipley, T. F., & Epstein, R. A. (2014). Variations in cognitive maps: Understanding individual differences in navigation. *Journal of Experimental Psychology: Learning Memory and Cognition*, *40*(3), 669–682. https://doi.org/10.1037/a0035261

Wernstedt, K., Roberts, P. S., Arvai, J., & Redmond, K. (2019). How emergency managers (mis?)interpret forecasts. *Disasters*, *43*(1), 88–109. https://doi.org/10.1111/disa.12293

Whyte, A. V., & Burton, I. (Eds.). (1980). *Environmental risk assessment*. John Wiley.

Wilson, R. S., Zwickle, A., & Walpole, H. (2019). Developing a Broadly Applicable Measure of Risk Perception. *Risk Analysis*, *39*(4), 777–791. https://doi.org/10.1111/risa.13207

Wisner, B., Blaikie, P., Cannon, T., & Davis, I. (2004). The challenge of disasters and our approach. In *At Risk: natural hazards, people's vulnerability and disasters* (Second, pp. 3–48). Routledge.

Witt, J. K., & Clegg, B. A. (2021). Dynamic ensemble visualizations to support understanding for uncertain trajectories. *Journal of Experimental Psychology: Applied*, *Advance on*. https://doi.org/10.1037/xap0000370

Yang, Z. J., Aloe, A. M., & Feeley, T. H. (2014). Risk Information Seeking and Processing Model: A Meta-Analysis. *Journal of Communication*, *64*, 20–41. https://doi.org/10.1111/jcom.12071

Yang, Z. J., & Kahlor, L. (2012). What , Me Worry? The Role of Affect in Information Seeking and Avoidance. *Science Communication*, *35*(2), 189–212. https://doi.org/10.1177/1075547012441873

Zack, E. S., Kennedy, J. M., & Long, J. S. (2019). Can nonprobability samples be used for social science research? A cautionary tale. *Survey Research Methods*, *13*(2), 215–227. https://doi.org/doi:10.18148/srm/2019.v13i2.7262 ?c

Zikmund-fisher, B. J., Smith, D. M., Ubel, P. A., & Fagerlin, A. (2007). Validation of the Subjective Numeracy Scale: Effects of Low Numeracy on Comprehension of Risk Communications and Utility Elicitations. *Medical Decision Making*, *27*(5), 663–671. https://doi.org/10.1177/0272989X07303824

APPENDIX

Subjective Numeracy Scale (SNS) (Fagerlin et al., 2007)

Cognitive abilities (1 = not at all good, 6 = extremely good)

1. How good are you at working with fractions?
2. How good are you at working with percentages?
3. How good are you at calculating a 15% tip?
4. How good are you at figuring out how much a shirt will cost if it is 25% off?

Preference for display of numeric information

5. When reading the newspaper, how helpful do you find tables and graphs that are parts of a story? (1 = not at all, 6 = extremely)
6. When people tell you the chance of something happening, do you prefer that they use words (‘‘it rarely happens’’) or numbers (‘‘there’s a 1% chance’’)? (1 = always prefer words, 6 = always prefer numbers)
7. When you hear a weather forecast, do you prefer predictions using percentages (e.g., ‘‘there will be a 20% chance of rain today’’) or predictions using only words (e.g., ‘‘there is a small chance of rain today’’)? (1 = always prefer percentages, 6 = always prefer words; reverse coded)
8. How often do you find numerical information to be useful? (1 = never, 6 = very often)

Santa Barbara Sense of Direction (SBSOD) Scale (Hegarty et al., 2002)

7-point Likert Scale (strongly disagree to strongly agree).

1. I am very good at giving directions.
2. I have a poor memory for where I left things. (REV)
3. I am very good at judging distances.
4. My “sense of direction” is very good.
5. I tend to think of my environment in terms of cardinal directions (N, S, E, W).
6. I very easily get lost in a new city. (REV)
7. I enjoy reading maps.
8. I have trouble understanding direction. (REV)
9. I am very good at reading maps.
10. I don’t remember routes very well while riding as a passenger in a car. (REV).
11. I don’t enjoy giving directions. (REV).
12. It’s not important to me to know where I am. (REV).
13. I usually let someone else do the navigational planning for long trips. (REV).
14. I can usually remember a new route after I have traveled it only once.
15. I don’t have a very good “mental map” of my environment. (REV)

Risk Perception Items (Wilson et al., 2019)

All items are measured on a 5-point Likert scale.

General

1. How risky is/are X?

Affect (i.e., concern/emotion)

2. How concerned are you (if at all) about X?
3. When you think about X for a moment, to what extent do you feel fearful?
4. When you think about X for a moment, to what extent do you feel anxious?
5. When you think about X for a moment, to what extent do you feel worried?
6. Considering any potential effects that X might have on you personally, how concerned are you about X?

Probability

7. How likely is it that X will occur/[you will do X] this year where you live?
8. I am confident that X will not occur/[I will not do X] this year where I live (REV)
9. How often do X occur where you live? [How often do you X?]

Consequences (i.e., severity)

10. If I did experience X, it is likely that it would negatively impact me.
11. If I did experience X, it would have a severe effect on me personally.

Hurricane Experience Score (HES) (Ehrlich et al., 2010)

Have you or has anyone in your household ever:

- evacuated or left your residence to go someplace safer in response to the threat of a hurricane?
- had damage to or loss of property because of a hurricane?
- had any other financial losses such as business losses or loss of income because of a hurricane?
- been injured (including loss of life) due to a hurricane?
- had emotional impacts or personal distress because of a hurricane?
    - Answered as (yes/no). Scored together, where more “yes” responses mean a higher experience score.

Interview Recruitment Email

Subject Line: Invite to Interview for Wind Exceedance Graphic Study

Dear [insert name],

My name is Zoey Rosen and I am a researcher from Colorado State University studying Environmental and Science Communication. I am working with the Cooperative Institute for Research in the Atmosphere (CIRA) at Colorado State University; its mission is to foster multi-disciplinary cooperation between NOAA research scientists and Colorado State University research staff, faculty, and students. I am interested in studying the perceptions of a new wind exceedance graphic product, risk perceptions surrounding wind, and decision-making with forecast graphics. You were recommended to me by XXX, and they think you would be a valuable addition to my study.

Your insights will contribute to an in-depth look into product design, risk perceptions, and decision-making with forecast graphic products. You are one of several emergency managers or meteorologists that I would like to interview for our research. The interview will take about an hour in-person, over the phone, or via video chat. Your identifying information, including name and organization, will not be connected to your data. Data will only be accessible on a password-protected and encrypted file. Data will be reported, in aggregate, in reports, academic papers, and/or presentations.

I aim to complete all of these interviews during September or October 2022. Would you be interested in participating? Please let me know, I would be happy to speak by phone or via email. If have any questions, please don't hesitate to ask.

Sincerely,

Zoey Rosen, M.S. | PhD Candidate | Colorado State University

Interview Oral Informed Consent Script

To begin, I'd like to thank you for your participation in my research interview. I'll be sharing some information regarding your rights as a participant in this research. As a reminder, my name is Zoey Rosen and I'm a graduate student at Colorado State University studying science communication. My principal investigator (PI) is Dr. Marilee Long (Marilee.long@colostate.edu; (970) 491-6463). My contact information is (zoey.rosen@colostate.edu; (818) 665-9383). My project focuses on studying the perceptions of a new wind exceedance graphic product, risk perceptions surrounding wind, and decision-making with forecast graphics.

Please understand that all of your correspondence with me is completely confidential and no identifying information will be attached to your data. This includes your name, demographics, specific location or any other potentially identifying information. Your data will be reported in sum with the other participants, but again, no identifying information will be included.

With your permission, I'd like to record this interview to refer to the data at later dates after all interviews are completed. This will help me to focus on our conversation in real time, instead of being focused on recording notes. This recording will be transcribed and only used for my analysis; the recording and transcription will be stored on a password protected computer and hard drive, and no one else will be able to access it. After data is analyzed, the recording and transcription will be destroyed. If you are uncomfortable at any time with the recorder, please let me know and I will turn it off.

Your cooperation and participation in this interview are completely voluntary. You can discontinue at any time with absolutely no risk or consequence. There are also no anticipated risks involved in this interview. Should you have any questions or concerns regarding your rights as a research participant, the Colorado State University Research Integrity and Compliance Review Office can be reached at 970-491-1553.

[Ask verbally and acquire response]

Do you understand what you are being asked to do?

Do you have any questions before we begin?

Do you agree to participate in the study?

May I use the audio recorder for our interview today?

Interview Guide

Before the interview (remotely) participants were asked to fill out the Likert scales on numeracy and spatial cognition. These scales are not anticipated to prime the answers in the interview. Note: The order of the questions in this interview guide do *not* follow the order of they hypotheses and research questions on purpose.

1. To start off, I'd like to know a little bit more about you.
    a. What is your occupation (e.g., EM or meteorologist)?
    b. How long have you been in this career?
    c. How much experience do you have with hurricanes personally? In a professional capacity?
2. Do you have a color-vision deficiency or consider yourself to be colorblind?
3. Show prototype graphics of past storms from the EM's/Met's local area (e.g., showing Ida to Gulf coast EMs, etc.)
    a. GRAYSCALE WITH POINT A AND/OR B
        i. What does the title of this graphic mean?
        ii. What is the worst-case scenario for windspeed at point A?
            1. How confident are you about your answer? (1-5)
        iii. What is the worst-case scenario for windspeed at point B?
            1. How confident are you about your answer? (1-5)
    b. GRAYSCALE
        i. What is the worst-case scenario for windspeed at [home location]?
            1. How confident are you about your answer?
    c. SHOW REDS-ONLY, YELLOW-TO-RED, AND TEXTURE

i. What do you think about the color schemes?

ii. Which graphic seems riskier? Why?

iii. Which version do you prefer? Why?

d. INTERSTATE/HIGHWAY OVERLAY ON COLOR PREFERENCE

i. What do you think about the interstate overlay?

ii. Does this version seem riskier than the previous version? Why or why not?

iii. Which version of the graphics you've seen so far do you prefer?

e. CITY LANDMARK ON COLOR PREFERENCE

i. What do you think about the city landmarks?

ii. Does this version seem riskier than the previous version? Why or why not?

iii. Which version of the graphics you've seen so far do you prefer?

f. What decisions could you make using the wind exceedance product?

MINIMIZE SHARE SCREEN FOR GENERAL CONVO

g. What risks are there for wind specifically from hurricanes?

h. What vulnerabilities exist in coastal communities?

i. Give definition of vulnerable (geographic, population, etc.)

4. Knowing the aim of my research, is there anything you'd like to reiterate from our conversation today?

a. Thank you again for your participation. Would you be available for follow-up questions for additional information? How would you prefer I contact you for a quick response?

Survey Informed Consent

Thank you for your interest in our hurricane graphic research! You must be at least 18 years of age or older to participate in this survey.

There are no known risks to participating in this survey. Through this survey, we hope to gain more knowledge about perceptions of a new hurricane graphic, risk perceptions surrounding wind, and decision-making with forecast graphics.

The survey is voluntary, anonymous, and should only take about 20 minutes of your time. If you decide to participate in the survey, you may withdraw your consent, stop the survey and exit at any time without penalty.

To indicate your consent to participate in this research, please click the consent button located below and proceed to the survey.

If you do not consent to the survey or are under the age of 18, please exit now.

If you have any questions about the research, please contact Zoey Rosen at zoey.rosen@colostate.edu. If you have any questions about your rights as a volunteer in this research, contact the Colorado State University Institutional Review Board at: RICRO_IRB@mail.colostate.edu; 970-491-1553.

◙ Yes, I have read the above procedures and information and consent to participate in the survey.

Other Demographic Data from Public Survey

Race/Ethnicity			
	Frequency	Percent of Responses	Percent of Sample
Spanish, Hispanic, or Latino	117	15.14%	18.75%
White/Caucasian	482	62.35%	77.24%
Black/African American	93	12.03%	14.90%
American Indian/Native American or Alaska Native	13	1.68%	2.08%
Asian	28	3.62%	4.49%
Native Hawaiian or Other Pacific Islander	6	0.78%	0.96%
Other	32	4.14%	5.13%
Prefer not to say	2	0.26%	0.32%

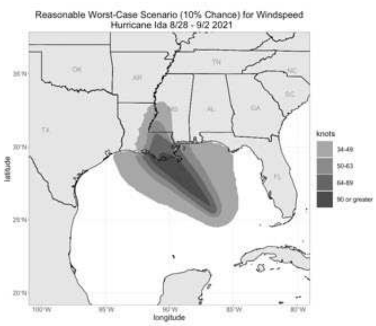
*773 selections (check all that apply). N = 624

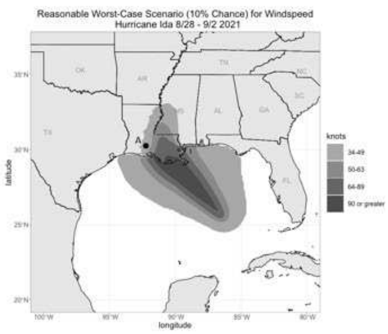
Highest Level of Education		
	Frequency	Percent
Some high school or less	21	3.40
High school diploma or GED	167	26.80
Some college, but no degree	128	20.50
Associates or technical degree	73	11.70
Bachelor's degree	149	23.90
Graduate or professional degree (MA, MS, MBA, MD, PhD, JD, DDS, etc.)	83	13.30
Prefer not to say	3	0.50
Total	624	100.00

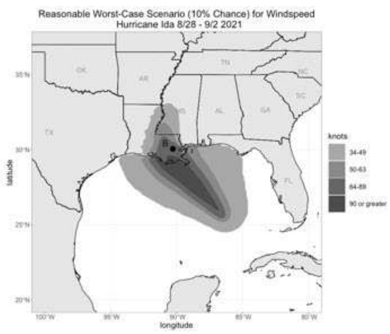
Employment Status		
	Frequency	Percent
Working-full time	281	45.00
Working part-time	84	13.50
Unemployed and looking for work	39	6.30
Homemaker or stay-at-home parent	25	4.00
Student	37	5.90
Retired	129	20.70
Other	29	4.60
Total	624	100.00

What was your total household income before taxes during the past 12 months?		
	Frequency	Percent
Less than $25,000	136	21.80
$25,000 - $49,000	133	21.30
$50,000 - $74,000	131	21.00
$75,000 - $99,000	83	13.30
$100,000 - $150,000	86	13.80
$150,000 or more	27	4.30
Prefer not to say	28	4.50
Total	624	100.00

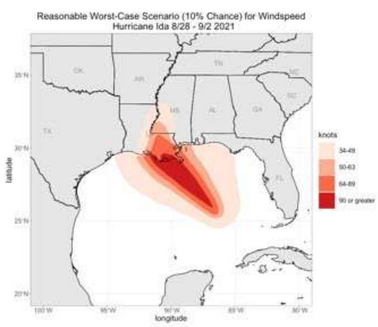
Prototype Graphics

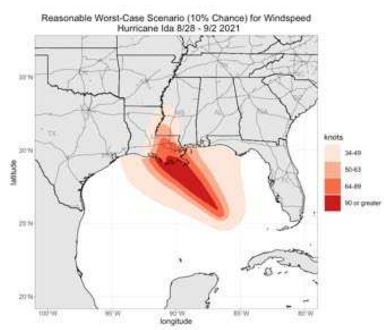


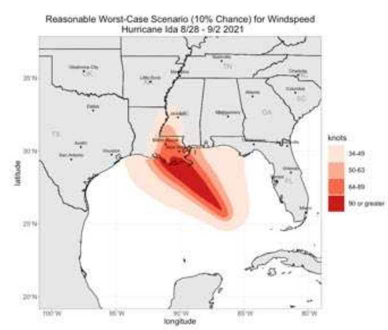




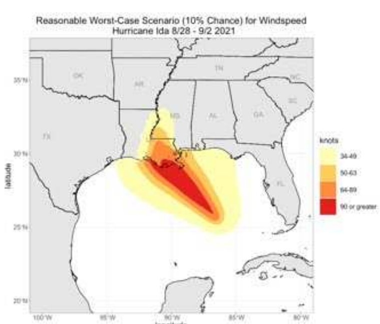
Base Graphic (Grayscale) Grayscale Point A (LA) Grayscale Point B (LA)

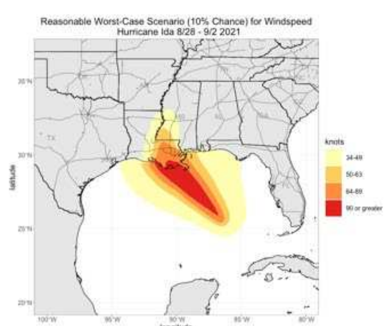


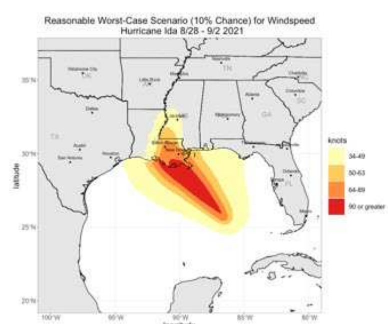




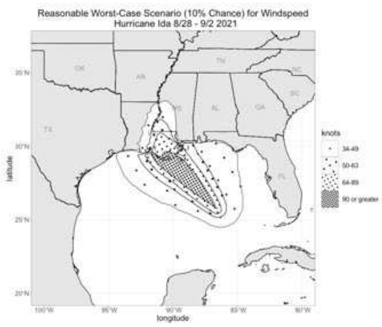
Reds-Only Graphic (LA) Reds-Only + Highways (LA) Reds-Only + Cities (LA)

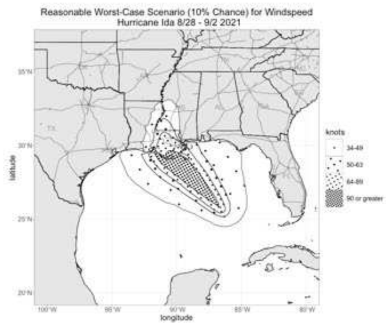


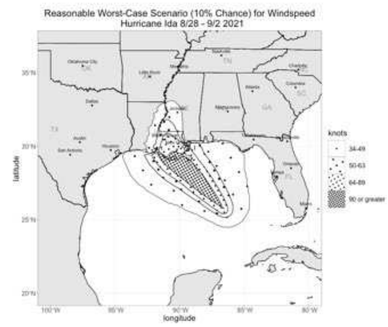




Yellow-to-Red Graphic (LA) Yellow-to-Red + Highways (LA) Yellow-to-Red + Cities (LA)



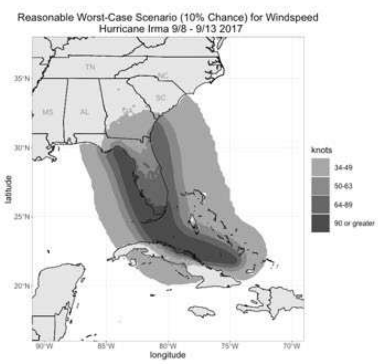


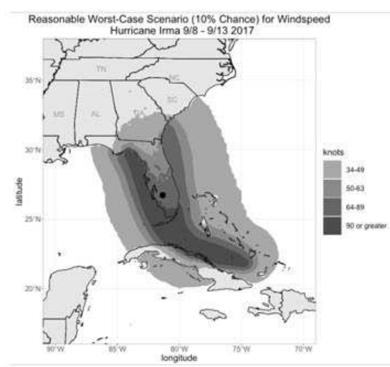


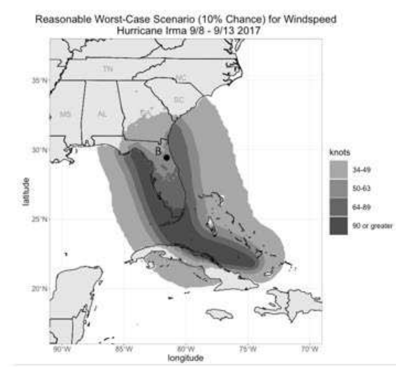
Texture Graphic (LA)

Texture + Highways (LA)

Texture + Cities (LA)



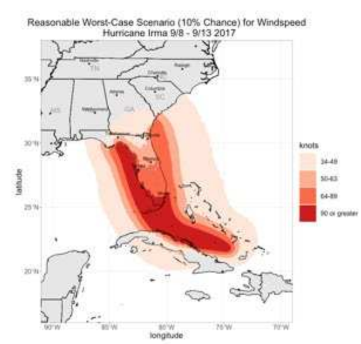




Base Graphic (Grayscale)

Grayscale Point A (FL)

Grayscale Point B (FL)



Reds-Only Graphic (FL)

Reds-Only + Highways (FL)

Reds-Only + Cities (FL)

Yellow-to-Red Graphic (FL)

Yellow-to-Red + Highways (FL)

Yellow-to-Red + Cities (FL)

Texture Graphic (FL)

Texture + Highways (FL)

Texture + Cities (FL)

Note: Survey for the public had wind in miles per hour (mph) instead of knots in the key, otherwise the graphics were the same.

Accuracy Measure Explanatory Figure

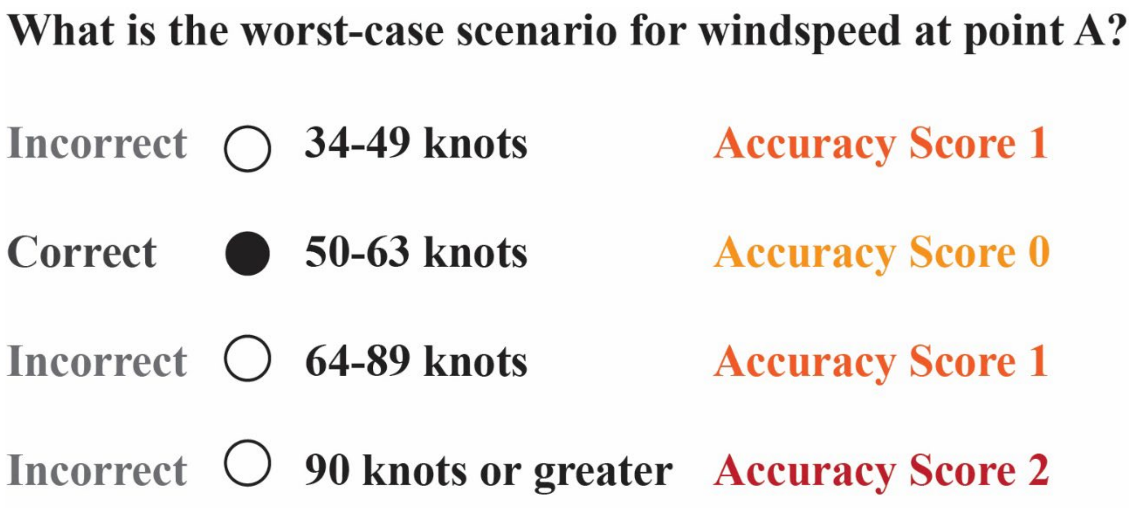


Figure 14. Visual diagram of how accuracy score was determined from participant responses