UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE VISUAL LANGUAGE OF TORNADO WARNINGS: PERCEPTION AND EMOTIONAL RESPONSE ACROSS FORMAT AND EXPOSURE

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF ARTS

By CODY FIELDS

Norman, Oklahoma

2025

THE VISUAL LANGUAGE OF TORNADO WARNINGS: PERCEPTION AND EMOTIONAL RESPONSE ACROSS FORMAT AND EXPOSURE

A THESIS APPROVED FOR THE

DEPARTMENT OF COMMUNICATION

BY THE COMMITTEE CONSISTING OF

Dr. Claude Miller, Chair

Dr. Elena Bessarabova

Dr. Mengfei Guan

Abstract

Tornadoes are one of the most dangerous weather hazards in the United States, prompting the National Weather Service to continuously evaluate how warnings are communicated to the public. Amid concerns that over-warning may erode public trust and a belief that current warnings can be further refined, this study employed a mixed design to test the effectiveness of traditional versus probabilistic tornado warning formats. A total of 347 participants were randomly assigned to receive one of the two formats across three simulated exposures. The study measured perceived likelihood, severity, message clarity, information sufficiency and seeking, and protective action intention, all components of the Protective Action Decision Model. It also measured cognitive load, Perceived Message Sensation Value, message fatigue, and desensitization, all novel applications from the field of health communication. While the probabilistic format was consistently rated as more novel, emotionally arousing, and dramatic—without invoking significant mental effort—traditional formats were rated significantly higher in perceptions of likelihood, severity, and protective action intention when averaged across all three exposures. Contrary to prevailing assumptions, participants did not perceive high message fatigue; in fact, individuals reporting higher fatigue indicated greater trust in forecasters. These findings suggest that probabilistic warnings amplify perceived threat in the most at-risk areas and reduce perceived threat in regions farther from the core impact zone. Fears of over-warning are not supported by the available data.

Keywords: Probabilistic tornado warning, message fatigue, cognitive load, Perceived Message Sensation Value, Protective Action Decision Model, desensitization

Table of Contents

The Visual Language of Tornado Warnings: Perception and Emotional Response Across Format and Exposure

On May 22, 2011, a large and destructive EF-5 tornado moved through downtown Joplin, Missouri, causing catastrophic damage and extensive loss of life. It tossed cars, leveled numerous homes, and damaged notable downtown structures such as the St. John's Regional Medical Center. With over 160 deaths, it became the single deadliest tornado dating back to 1950. Total damage topped $2.8 billion, cementing the storm as the costliest tornado on record in United States history (KY3 Staff, 2022). The Joplin tornado was a shocking outlier in modern history given consistent progress toward more accurate forecasts and longer warning lead times that have generally decreased the number of fatal tornadoes each year (Schmid, 2011).

The devastation wrought by the tornado highlights the work the National Weather Service (NWS) must still conduct in improving preparedness for violent tornadoes. In advance of conditions that favor the formation of tornadoes, the Storm Prediction Center in Norman, Oklahoma, issues tornado watches for broad geographic locations. As thunderstorms develop and the formation of tornadoes appears imminent, local NWS offices issue tornado warnings National Weather Service, n.d.). Between 2016 and 2020, the average time between the issuance of a tornado warning and the touchdown of a tornado varied between 8 to 10 minutes. However, approximately 70 percent of tornado warnings issued by the NWS constitute false alarms, or warnings that did not include a confirmed tornado (Erdman, 2021). Extensive debate is currently occurring among meteorologists regarding the importance of focusing on over-warning. A generous approach to warning may result in fatigue and decreased protective action, whereas a conservative approach to warning may result in missed tornadoes and increased loss of life (Brooks & Correia, 2018).

The NWS's mission is to protect life and property, and as such, the organization continues to develop new ways to refine their approach to saving lives. Accordingly, social scientists are currently working in conjunction with forecasters at the NWS to test a new way to visualize tornado warnings. The project has previously been referred to as Forecasting a Continuum of Environmental Threats, but more recently, it is known as Threats in Motion. This project seeks to revolutionize tornado warning information by utilizing probabilistic displays, where lighter colors/lower probabilities suggest reduced risk and darker colors/higher probabilities indicate increased risk. In addition, scientists are investigating whether the current weather software can support dynamically updating warnings in real time, representing a major shift from the static warning polygons currently in use (Stumpf & Gerard, 2021).

While some literature has explored how various tornado warning displays affect perception, the resultant findings have been inconsistent. Furthermore, some studies have suffered from the lack of an established theoretical framework, as well as a restrictive sample size. Further, almost none of these studies explore the cognitive or emotional responses to varying tornado warning formats. Thus, this paper seeks to contribute additional information to probabilistic warning information across a broader sample size while simultaneously providing novel information about the mental strain and emotions that these warnings elicit from individuals. In doing so, it will provide additional data for the NWS to examine in their quest to refine warning messages and protect citizens.

Literature Review

Theoretical Framework

Lindell and Perry (2012) devised the Protective Action Decision Model (PADM) to highlight the various factors that influence individual decision making during a crisis. This model begins with a series of factors that reside under the umbrella term of environmental and social context, including environmental cues, social cues, information sources, channel access and preference, warning messages, and receiver characteristics. These factors initiate pre-decisional processes, such as exposure, attention, and comprehension. These processes, in turn, shape individuals' perceptions of environmental threats, protective actions, and the credibility of stakeholders (Lindell and Perry, 2012). All these factors ultimately contribute to protective action decision making, in which individuals draw on prior experiences, evaluate the effectiveness of potential actions, consider whether additional information is needed, seek out that information, and ultimately take steps to protect themselves from the threat (Lindell and Perry, 2012).

Having navigated the pre-decisional processes, assessed perceptions of threat, protective actions, and stakeholders, and entered protective action decision making, the PADM suggests that individuals continue toward behavioral response. This includes information searching, protective action, and emotion-focused coping. This step is complicated by competing situation factors that may either assist or inhibit an individual from acting against a threat. Additionally, the model suggests a feedback loop originating from behavioral response that returns individuals back to the beginning environment and social context stage, usually because individuals seek to either validate or invalidate prior warnings they have received, as well as seek out additional sources of information to rely upon (Lindell & Perry, 2012).

Consistent with this framework, the present study operationalized threat perceptions specifically through perceived likelihood and perceived severity, reflecting participants' assessment of both the probability of a tornado event occurring and the anticipated consequences of that event.

PADM Application

Traditional Tornado Warnings

The PADM has been applied to a variety of different meteorological contexts, including tornadoes, hurricanes, and wildfires. As it relates to tornadoes, prior studies have devoted a significant amount of research on people's responses to traditional tornado warning polygons. The literature suggests that traditional tornado warning polygons strongly promote the tendency for people to perceive the centroid of the warning as the location of great risk, particularly when no contextual information is given to the graphic display (Ash et al., 2014; Jon et al., 2019; Lindell et al., 2016; Sherman-Morris, 2005). Consistent with this finding, Jon et al. (2018, 2019) found that individuals fail to recognize there is a much greater risk of danger along the near edge of the polygon versus the centroid. Although it is the NWS's intention that individuals perceive a significantly greater degree of risk just inside the polygon versus just outside the polygon, studies suggest only a slight polygon edge effect in perception of risk (Ash et al., 2014).

Additional research found that higher perceptions of being in the tornado's path were associated with decreased expectations of resuming normal activities and increased intention to seek information, shelter, or evacuate (Ash et al., 2014; Jon et al., 2018, 2019; Lindell et al., 2016). Differences in response appear to vary by gender, with females being less likely to ignore warning messages and more likely to take protective actions (Lindell, 2020). However, no other demographic variables exhibit a statistically significant relationship between perceptions of risk and expected response actions (Jon et al., 2018, 2019; Lindell et al., 2016; Miran et al., 2020).

This is consistent with similar research for hurricane evacuations (Huang et al., 2016). Taken together, this literature suggests that traditional tornado warning displays erroneously promote the perception that the greatest risk is at the center of the warning. This results in individuals closer to the tornado perceiving less danger, and individuals farther from the tornado perceiving greater danger, than is truly the case. One potential remedy to this predicament is the utilization of probabilistic displays, which would highlight the greatest threat closer to the tornado.

Probabilistic Tornado Warnings

Rather than a solid-colored warning polygon, probabilistic tornado warnings utilize a probability matrix to convey the odds of a particular location being affected by a tornado. Typically, lighter colors are utilized for lower probabilities farther from the tornado path while darker colors are utilized for higher probabilities closer to the tornado. Only a handful of studies have sought to explore how the public perceives this novel form of tornado warning, to mixed result. Some studies have found that probabilistic tornado warning information can broadly increase risk perception, cognition, and protective action (Klockow, 2013; Miran et al., 2017), while others have found that these outcomes are highly dependent on the tornado likelihood threshold (Miran et al., 2018; Miran et al., 2019; Miran et al., 2020; Qin et al., 2024). Some research was unable to establish a consistent advantage in probabilistic displays, with variance depending on the exact design and color scheme utilized (Ash et al., 2014).

Utilizing the literature previously discussed, the following hypotheses are presented:

H1a: Probabilistic tornado warnings are associated with greater perceived likelihood of tornado impact relative to traditional tornado warnings.

H1b: Probabilistic tornado warnings are associated with greater perceived severity of tornado impact relative to traditional tornado warnings.

H1c: Probabilistic tornado warnings are associated with greater perceived information sufficiency relative to traditional tornado warnings.

H1d: Probabilistic tornado warnings are associated with lower intentions to seek additional sources of information relative to traditional tornado warnings.

H1e: Probabilistic tornado warnings are associated with greater intentions to seek protective action relative to traditional tornado warnings.

Yet, despite the potential advantages that probabilistic warnings may bring, there are important limitations to keep in mind. For example, previous literature suggests a reluctance among scientists to utilize probabilistic displays when conveying information to a general audience, namely because they fear the uncertainty information may overwhelm or confuse viewers (Hullman, 2020). This underscores the need to explore whether probabilistic tornado warning displays, while increasing perceived risk, may ultimately prove inferior to traditional warning displays because they are mentally taxing or even delay the decision-making process during times when every second counts.

Cognitive Load

Grounded in educational psychology, the theory of cognitive load was first introduced in the 1980s through observing learners as they completed problems (Sweller, 1988). According to the theory, individuals learn through a mixture of working memory (short-term) and long-term memory processes (Cowan, 2008). Long-term memory stores information in a way that can be easily retrievable to solve problems over time (Paas et al., 2003). By contrast, the working memory system is constrained in that an individual can only process a limited amount of information simultaneously without losing it over time (Jonides et al., 2005). Given this fact, cognitive load encompasses two elements: mental load and mental effort. Mental load refers to concepts such as the complexity and presentation of a problem. As the name suggests, mental effort refers to the cognitive work necessary to learn and complete a task (Paas 1992; Paas et al., 2003). Notably, Sweller et al. (2019) point out that, "Cognitive load is increased when unnecessary demands are imposed on the cognitive system...[c]ognitive load may also be increased by processes that are germane to learning, such as instructional methods that emphasize subject information that is intrinsically complex" (p. 262).

Though not part of the original theory, recent literature concerning cognitive load has introduced three subcategories within the framework, all which compound to the overall cognitive load: intrinsic load, extraneous load, and germane load (Sweller et al., 1998). Intrinsic load relates to a learning task's complexity (Klepsch et al., 2017) while extraneous load refers to the presentation of the learning materials (Sweller et al., 2019). Germane load differs from the previous two in that it is not inherently negative. Rather, it describes how much effort a person chooses to spend to meaningfully understand learning material (Klepsch et al., 2017). Holistically, cognitive load has been found to contribute to information overload, reducing learning efficiency, reducing retention of information, and perhaps most importantly, constraining the decision-making process (Firth et al., 2019).

Because cognitive load is an education psychology term, there is little research extending the concept into a meteorological context. Some research has established that immersive virtual reality in classrooms or the redesign of online tools may reduce cognitive load and enhance the understanding of weather- and climate-related topics (Annetta & Newton, 2023; Calvo et al., 2022; Liu et al., 2022). Other studies have explored how cognitive load influences the usability of smartphone weather applications during times of severe weather (Khamaj, 2019). Aviation-focused literature found that the symbology and colors used during weather presentations for pilots significantly influenced their cognitive load and understand (Ahlstrom, 2005; Ahlstrom & Dworsky, 2012). Perhaps most relevant to this study is a 2013 journal article which found that complex graphical displays hinder an individual's ability to make optimal decisions (Fox et al., 2013).

Utilizing the literature previously discussed, the following hypotheses are presented:

H2a: Probabilistic tornado warnings are associated with greater cognitive load relative to traditional tornado warnings.

H2b: Higher cognitive load is negatively associated with intentions to seek protective action.

While warning format and cognitive load reflect how individuals process tornado warnings cognitively, it is equally important to consider their emotional and sensory reactions. This shift in focus is captured by Perceived Message Sensation Value (PMSV), which emphasizes how message features shape audience engagement.

Perceived Message Sensation Value

PMSV refers to how aspects of a message elicit sensory, affective, and arousal responses in an individual (Everett & Palmgreen, 1995). Early research predating PMSV established that individuals require a certain level of stimulation to devote attention to, and be persuaded by, a message (Harrington et al., 2003; Morgan et al., 2003). Subsequent literature established the idea of high sensation seekers or low sensation seekers, depending on the level of arousal that a message provides (Donohew et al., 1998; Zuckerman, 1988). In the principal PMSV study, the authors developed a 17-item scale that included four elements: emotional impact, physiological impact, sensory impact, and novelty (Everett & Palmgreen, 1995). However, owing to a limited sample size that prevented a factor analysis, additional research refined this scale to establish three important factors: emotional arousal, dramatic impact, and novelty (Stephenson & Palmgreen, 2001).

In other words, messages high in sensation value are novel, creative, or unusual; complex; involve stimuli that are emotionally powerful or arousing; graphic or explicit; somewhat ambiguous; unconventional; fast-paced; and suspenseful (Everett & Palmgreen, 1995; Palmgreen & Donohew, 2003; Noar et al., 2010). While all these features need not be present to elicit greater attention from an individual, the effects are additive, particularly as it relates to novelty (Palmgreen & Donohew, 2003). Throughout the 2000s, literature drew a distinction between PSMV and message sensation value (MSV). While both concepts directly relate to the principles of emotional arousal, dramatic impact, and novelty, PMSV measures the subjective evaluation of a message by participants, whereas MSV is operationalized as an objective evaluation of a message by analyzing the formal video, audio, and content dimensions (Morgan et al., 2003).

Within a health communication context, PMSV has been directly related to the persuasiveness of a message. Specifically, research utilizing self-reported measures has established that visual graphics (Thorson & Lang, 1992) and emotionally arousing messages (Lang et al., 1996) affect arousal and cognitive outcomes. Empirical studies tracking measures such as heart rate and skin conductance have reaffirmed these findings. Additionally, the effectiveness of the message seems equally strong for high sensation seekers and low sensation seekers (Xu, 2015). Unfortunately, because PMSV has been applied largely to health communication, very little literature exists within a meteorological context. A sole study has explored how the high PMSV of a climate change film heightened the risk perception among viewers (Lin, 2013). Otherwise, there is a clear avenue to extend research about PMSV into an entirely new arena.

Utilizing the literature previously discussed, the following hypotheses are presented:

H3a: Probabilistic tornado warnings are associated with greater emotional arousal relative to traditional tornado warnings.

H3b: Probabilistic tornado warnings are associated with greater dramatic impact relative to traditional tornado warnings.

H3c: Probabilistic tornado warnings are associated with greater novelty relative to traditional tornado warnings.

H3d: Probabilistic tornado warnings are associated with greater accuracy relative to traditional tornado warnings.

H3e: Probabilistic tornado warnings are associated with a greater sense of urgency to act relative to traditional tornado warnings.

While PMSV highlights the importance of emotionally engaging and visually impactful warning messages, a common concern among meteorologists is that over-warning, regardless of the warning display, will lead to a disengaged public. In the context of tornado warnings, where false alarms and frequent alerts are common—particularly in the Midwest and Southern United

States—understanding how message fatigue interacts with an individual's response is critical to explore.

Message Fatigue and Desensitization

Message fatigue refers to feelings of exhaustion, boredom, and decreased motivation resulting from prolonged exposure to redundant health messages. So et al. (2017) highlighted research from prior studies which found that message avoidance, annoyance, information seeking, desensitization, cognitive elaboration, and counterargument are all key variables in quantifying message fatigue. The study sought to build upon this literature by proposing, and validating, four dimensions of message fatigue. The first dimension included perceived overexposure, or an individual's belief they received a health-related message too frequently. The second dimension referenced perceived redundancy, or an individual's belief that messages fail to provide new information. The third dimension referenced exhaustion, or an individual's feeling of being overwhelmed by the messages. The fourth and final dimension referenced tedium, or an individual's feeling or boredom from the messages.

Meanwhile, Cho & Salmon (2007) found that message fatigue may, over time, lead to an emotional response known as desensitization. Drawing upon previous literature, the authors noted that repeated exposure to health messages may render the public apathetic over the long run. This concept was first affirmed by Lazarsfeld and Merton (1948) which noted that mass media saturates the public with information, reducing their willingness to act politically on social issues. Subsequent research argued that even powerful images, such as photos of dead soldiers during times of war, lose their impact after they have been viewed repetitively (Downs, 1972). More recently, Kinnick et al. (1996) extended this research to demonstrate that long-term media coverage of crime, homeless, AIDS, and child abuse resulted in desensitization toward these issues among the public.

A widespread concern within the meteorological community is the fear that issuing too many tornado warnings may result in fatigue or complete disregard of the warning message. Many of these fears appear to originate from a singular event, the Joplin tornado. While literature published after that event found that frequent false alarms may indeed influence whether people respond to the tornado warning out of overexposure (Paul et al., 2015), very few studies have been conducted to validate this fear. A principal study in 2019 found that concerns about false alarms leading to a complacent public are overblown, namely because individuals assess tornado warnings to be more accurate than they truly are. The percentage of false alarms and the number of tornado warnings did not predict taking protective action (Lim et al., 2019). Otherwise, little literature exists to apply these concepts in relation to tornadoes. Some application exists in relation to wildfires, with mixed results in whether overexposure to warning messages impacts willingness to act (Hanson-Easey et al., 2019; Islam et al., 2024; Mackie, 2014). Meanwhile, an examination of Bangladesh tropical cyclones found that overexposure to failed warning messages inhibited protective action (Ahsan et al., 2016).

There is a notable gap in research on message fatigue and desensitization in meteorological contexts. Repetitive tornado warnings may parallel repetitive health messages, where studies show fatigue can lead to serious outcomes, including the adoption of message-inconsistent beliefs and reduced protective actions (Keating & Skurka, 2024). Considering the potentially catastrophic consequences of extreme weather events—such as tropical cyclones, flash floods, wildfires, or tornadoes—the effects of message fatigue and desensitization in this context could have life-threatening implications. Moreover, further research into message fatigue and desensitization could help quantify forecasters' anecdotal concerns that excessive warnings may pose a significant risk to public safety.

Utilizing the literature previously discussed, the following hypotheses are presented:

H4a: Greater message fatigue is positively associated with lower trust in forecasters due to repeated false alarms.

H4b: Message fatigue moderates the effect of warning format on perceived likelihood, such that the effect of format is weaker for individuals with higher fatigue.

H4c: Message fatigue moderates the effect of warning format on PMSV outcomes (emotional arousal, dramatic impact, novelty), such that individuals with higher fatigue report lower PMSV regardless of format.

H4d: Probabilistic tornado warnings are associated with lower desensitization relative to traditional tornado warnings.

Methods

Participants

A total of 347 participants were recruited for this study, including 72% (n = 250) via CloudResearch and 28% (n = 97) via the University of Oklahoma's research opportunities pool (SONA). Participants ranged from 18 to 75 years (M = 32.9, SD = 12.8, n = 332). The sample was 48.7% male (n = 169), 49.3% female (n = 171), and 1.7% non-binary/third gender (n = 6); 0.3% (n = 1) preferred not to state their gender. The sample was 75.2% (n = 261) White, 10.4% (n = 36) Black or African American, 5.8% (n = 20) Hispanic or Latino, 2.3% (n = 8) American Indian or Alaska Native, 4.6% (n = 16) Asian, 0.3% (n = 1) Native Hawaiian or Pacific Islander, and 0.3% (*n* = 1) Other. In terms of educational attainment, 0.6% (*n* = 2) had not completed high school, 13.8% (*n* = 48) held a high school diploma or GED, 38.9% (*n* = 135) had some college experience, 11.0% (*n* = 38) held an associate degree, 24.5% (*n* = 85) held a bachelor's degree, 8.4% (*n* = 29) held a master's degree, and 2.9% (*n* = 10) held a doctorate. Collectively, 23.1% of participants (*n* = 80) lived in an independently owned or rented residence while 76.9% (*n* = 267) did not. Additionally, 28% (*n* = 97) reported having children while 72% (*n* = 249) did not.

The geographic distribution of participants was divided into two regions: the Midwest and Southern regions of the United States. The Midwest region included the states of Iowa, Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin. The Southern region included the states of Alabama, Arkansas, the District of Columbia, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia. Both regions were defined using CloudResearch's US Regions tool. Of the 344 participants, 74% (*n* = 255) resided in the Southern region while 26% (*n* = 89) resided in the Midwest region. Since the study sought individuals with prior tornado experience, participation outside these regions was restricted via CloudResearch, and SONA access was necessarily limited to University of Oklahoma students. Furthermore, only participants that spoke English were recruited for this study.

In all, 333 participants (96%) reported prior experience with either a tornado warning or a confirmed tornado, while only 14 (4%) did not. Asked to estimate exposure to tornado warnings, 52 participants (15%) responded they receive tornado warnings rarely (less than one per year), 132 (38%) responded occasionally (one to two times per year), 88 (25.4%) responded frequently (three to five times per year), and 75 (21.6%) responded very frequently (more than five times per year). Asked to rate their familiarity with the tornado warning system, including the design of the warnings and the messages that accompany them, 9 participants (2.6%) were very unfamiliar, 12 (3.5%) were unfamiliar, 30 (8.6%) were somewhat unfamiliar, 36 (10.4%) were neutral, 103 (29.7%) were somewhat familiar, 99 (28.5%) were familiar, and 58 (16.7%) were very familiar.

Procedure

This study utilized a 2 (Warning Format: traditional vs. probabilistic) × 3 (Exposure: first, second, third warning location) mixed factorial design. The between-subjects factor was the format of the tornado warning, in which participants were randomly assigned to view either a traditional warning or a probabilistic warning. Participants were assigned into one of these two conditions by Qualtrics' built-in randomizer function to ensure a roughly equal number of participants per group. The within-subjects factor was the warning exposures, in which all participants of the study viewed three sequential tornado warnings, each depicting their location in a different part of the polygon (near the beginning, middle, and end of the polygon).

After providing informed consent, participants answered basic demographic questions, including their age, gender, ethnicity, highest level of education, state of residence, living situation, and whether they had children. Subsequent questions assessed participants' prior tornado experience, including whether they had personally experienced a tornado or a tornado warning. If so, they were asked whether they had been injured, sustained property damage, taken shelter during a warning, or chosen to ignore one. Additional items measured the frequency of tornado warnings in their area per year and their familiarity with the current tornado warning system. After participants were exposed to a message fatigue scale, they were evenly divided into two conditions. The traditional warning format condition included a graphical display featuring a solid red polygon consistent with operational tornado warnings. The probabilistic tornado warning format condition included a graphical display featuring a color-coded probability polygon, with gradient colors representing the odds of a given location being impacted by the tornado. Each participant viewed three warnings within their assigned condition, with the participant's location varying across the warnings: at the beginning, the middle, and the end of the warning area (Figure 1). The order of presentation was not randomized. After each individual exposure, they completed a series of questions related to perceived likelihood, severity, message clarity, information sufficiency and seeking, protective action intention, and cognitive load. After all three warnings, they answered questions measuring PMSV and desensitization.

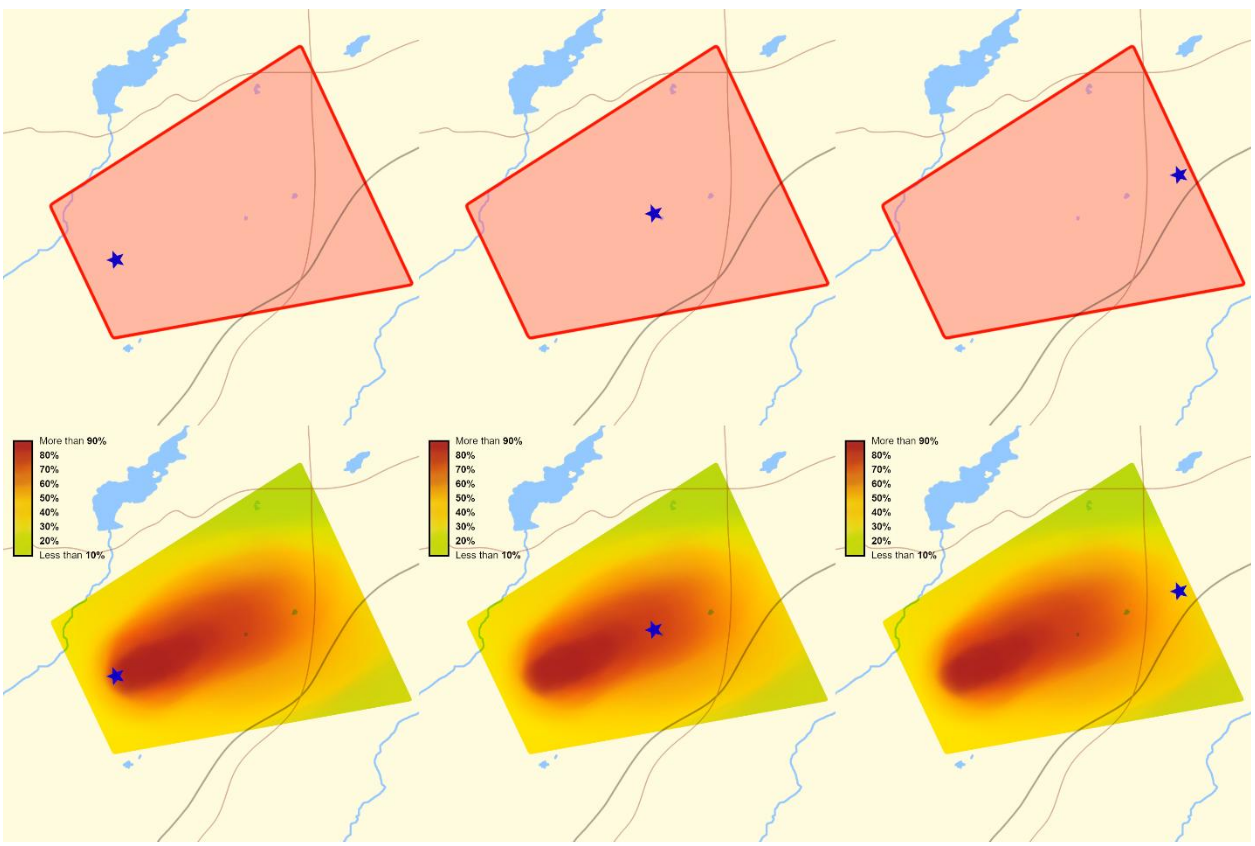

Figure 1: Tornado warning displays used during the study

To test my study hypotheses, a series of statistical analyses were employed after restructuring the data into a long format. H1a through H2a were examined using mixed-effects repeated-measures models employing an autoregressive covariance structure that assessed the main effects of warning format, exposure, and their interaction on key dependent variables: perceived likelihood (H1a), perceived severity (H1b), information sufficiency (H1c), information seeking (H1d), protective action intention (H1e), and cognitive load (H2a). In all cases, the subject identifier was set as the random effect to account for individual variability. Exposure was entered as a within-subjects factor, and warning format was entered as a between-subjects factor. Interactions between exposure, warning format, and demographic variables were included to model the dynamics over time. Meanwhile, H2b, predicting that higher cognitive load would be associated with lower protective action intention, was evaluated using Pearson's correlation.

The dimensions of PMSV—emotional arousal, dramatic impact, and novelty (H3a-H3c)—were assessed with a multivariate analysis of variance (MANOVA) followed by subsequent univariate tests. H3d and H3e, which tested differences in perceived accuracy and urgency between warning formats, were examined using independent samples t-tests. H4a, concerning the relationship between message fatigue and trust in forecasters, was also assessed via Pearson's correlation. Finally, hierarchical regression analyses were conducted to evaluate whether message fatigue moderated the effects of warning format on perceived likelihood (H4b) and PMSV outcomes (H4c), while H4d—predicting that desensitization would differ between warning formats—was examined using an independent samples t-test.

This study was approved by the University of Oklahoma's Institutional Review Board (IRB #18342), and all participants provided informed consent prior to beginning the survey. All statistical analyses were conducted using IBM SPSS Statistics Version 29. Three attention check items were embedded throughout the survey: one was presented to all participants while the other two were condition specific. Participants were excluded from the final dataset if they failed both the universal attention check and the attention check assigned to their condition.

Additionally, participants who completed only a small portion of the survey or skipped a majority of any one principal section (message fatigue, PADM, cognitive load, PMSV, or desensitization) were removed from analysis. After three exclusions, the final analytic sample included 347 participants.

Measures

Message Fatigue

Message fatigue was measured using a 17-item scale, with each item rated on a 7-point Likert scale ranging from *strongly disagree* to *strongly agree*. The scale was adapted from Williams (2020), which itself was based on So et al. (2017). Items assessed emotional and cognitive weariness resulting from repeated tornado warning exposure, including feelings of annoyance, frustration, disengagement, and reduced perceived urgency. An example item is, "At this point, I've heard about threats related to tornadoes more than I ever needed to."

Protective Action Decision Model

Constructs of the PADM included perceived likelihood, perceived severity, message clarity, information sufficiency, information seeking, and protective action intention. Perceived likelihood was measured with the item, "How likely is a tornado to impact your specific location based on this warning?" utilizing a 7-point Likert scale ranging from *extremely unlikely* to *extremely likely*. Perceived severity was measured with the item, "How severe do you believe the effects of this tornado could be for your area?" using a 5-point Likert scale ranging from *not severe* to *extremely severe*. Although we considered creating a construct that combined perceived likelihood and severity under the umbrella of threat perception, these constructs were only moderately correlated, $r(344) = .56$, $p < .001$, indicating related but distinct components.

Therefore, they were analyzed as separate variables. Message clarity was measured using the statement, “This tornado warning clearly explains what actions I should take and makes those actions seem easy to implement.” Information sufficiency was measured with the item, “This tornado warning provides both accurate and sufficient information for me to assess the risk to my safety,” Information seeking was measured with the item, “This warning makes me want to seek additional information and verify the risk with other sources.” All aforementioned items were measured using a 7-point Likert scale ranging from *strongly disagree* to *strongly agree*. A follow-up question asked participants to check which sources they would trust most for information, with options including television, radio, social media, internet websites, weather apps, text alerts from authorities, and friends or family members. Protective action intention was measured with the item, “Would you decide to take immediate action after seeing this warning?” using a 5-point Likert scale ranging from *definitely not* to *definitely yes*.

Cognitive Load

Cognitive load was measured using a single-item scale adopted from Paas (1992). The item asked, “Please rate the amount of mental effort you invested in interpreting this tornado warning” using a 9-point Likert scale ranging from *very, very low mental effort* to *very, very high mental effort*. Although subsequent studies have developed larger-item scales, the original Paas scale continues to be the most popular method of measuring cognitive load among researchers (Sweller, 2018).

Perceived Message Sensation Value

A novel scale was developed to measure PMSV. Because it is typically used to assess the audio-visual elements of public service announcements in health communication campaigns, many commonly used items were not appropriate for application to tornado warning graphics. Instead, a new scale was created based on the core dimensions of PMSV—emotional arousal, dramatic impact, and novelty—as well as their potential influence on source trust, accuracy, and behavioral intention. Emotional arousal was measured with the item, “The colors and design of these warnings made the tornado threat feel real and immediate.” Dramatic impact was measured with the item, “The way these tornado warnings were visually displayed lacked intensity and made the situation feel less serious.” Novelty was measured with the item, “The way these tornado warnings were visually presented felt fresh and different from other tornado warnings I have seen.” Source trust was measured with the item, “When tornado warnings repeatedly fail to correspond with actual tornado events, I begin to trust forecasters less.” Accuracy was measured with the item, “The way these tornado warnings were visually displayed made me question their accuracy.” Behavioral intention was measured with the item, “The way this tornado warning was visually displayed made me feel less urgency to take shelter.” In all cases, items were measured using a 7-point Likert scale ranging from *strongly disagree* to *strongly agree*.

Desensitization

Desensitization was measured using a 4-item scale adapted from So et al. (2017). The four items included, “Tornado-related messages no longer arouse any emotions in me,” “I am immune (not responsive) to tornado-related messages,” “I feel indifferent about tornado-related messages,” and “I feel numb to tornado-related messages.” All four items were measured using a 7-point Likert scale ranging from *strongly disagree* to *strongly agree*.

Results

Protective Action Decision Model

Mixed-effects repeated-measures models were conducted to examine whether perceived likelihood, severity, message clarity, information sufficiency and seeking, and protective action intentions varied as a function of warning format (traditional vs. probabilistic), exposure order (first, second, third warning), and their interaction. Detailed statistics are provided in Table 1. Unless specifically referenced, no other demographic variables significantly moderated the effects of warning format, exposure, or their interaction on any of the PADM variables (all $p >$ .05).

PADM LMM Results

Table 1: Linear Mixed Model Tests of Fixed Effects for PADM Outcomes

DV	Effect	df (Num, Den)	F	p
Likelihood	Format	1, 338.6	4.81	.029
	Exposure	2, 643.7	329.51	<.001
	Format × Exposure	2, 643.7	69.90	<.001
Severity	Format	1, 359.4	4.37	.037
	Exposure	2, 668.0	325.66	<.001
	Format × Exposure	2, 668.0	101.93	<.001
Message Clarity	Format	1, 361.6	3.57	.060
	Exposure	2, 681.6	33.71	<.001
	Format × Exposure	2, 681.6	0.96	.382
Information Sufficiency	Format	1, 355.4	0.01	.924
	Exposure	2, 677.7	13.48	<.001
	Format × Exposure	2, 677.7	7.05	<.001
Information Seeking	Format	1, 357.5	0.05	.823
	Exposure	2, 679.1	20.63	<.001
	Format × Exposure	2, 679.1	0.01	.987
Protective Action	Format	1, 345.4	16.45	<.001
	Exposure	2, 661.9	231.17	<.001
	Format × Exposure	2, 661.9	44.58	<.001

Table 1: Tests of fixed effects for PADM variables

Likelihood

Results indicated a significant main effect of warning format on perceived likelihood, with probabilistic warnings ($M = 5.25$, $SD = .76$) generating slightly lower perceptions of likelihood compared to traditional warnings ($M = 5.49$, $SD = 1.08$). Therefore, H1a's assumption that participants in the probabilistic group would report greater perceived likelihood was not supported by the available data. There was also a significant main effect of exposure. Perceived likelihood decreased steadily from Exposure 1 ($M = 6.03$, $SD = 1.36$) to Exposure 3 ($M = 4.11$, $SD = 1.60$). Further, a significant interaction emerged between warning format and exposure (Fig. 2). Follow-up estimated marginal means revealed that probabilistic warnings generated higher perceived likelihood at Exposure 1 ($M = 6.50$, $SD = 1.18$) compared to traditional warnings ($M = 5.58$, $SD = 1.37$), but this pattern reversed by Exposure 3, where traditional warnings elicited higher likelihood perceptions ($M = 4.7$, $SD = 1.49$) than probabilistic warnings ($M = 3.50$, $SD = 1.48$). Pairwise comparisons with Bonferroni adjustment indicated that all exposure contrasts were significant at $p < .001$, with perceived likelihood decreasing across each subsequent exposure for both warning formats.

Education moderated the combined effects of warning format and exposure, such that the pattern of likelihood perceptions over time varied slightly by education level. Similarly, having children moderated the warning format by exposure interaction, suggesting differing patterns between parents and non-parents. Warning frequency moderated the effect of exposure alone, indicating variation in likelihood perceptions across exposures depending on how often participants typically receive warnings.

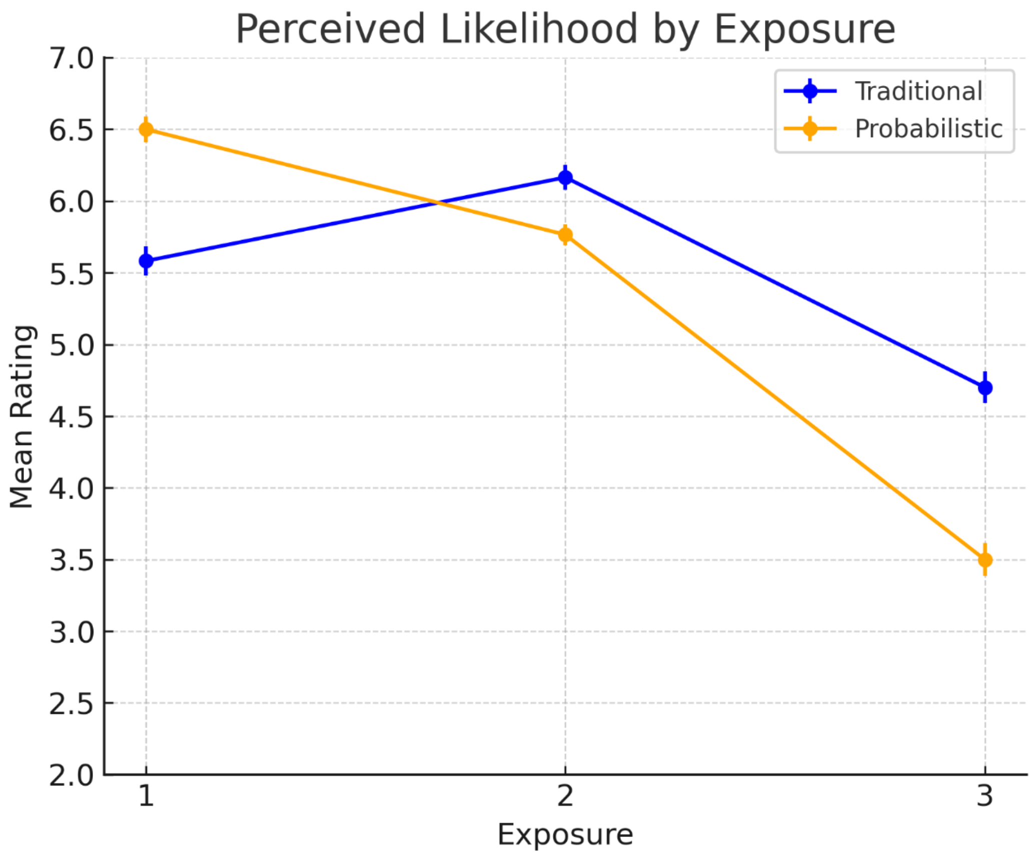

Figure 2: Perceived likelihood across warning format by exposure

Severity

Results revealed a significant main effect of warning format on perceived severity. Participants in the traditional warning condition reported higher perceived severity ($M = 3.51$, $SD = .84$) compared to those in the probabilistic warning condition ($M = 3.35$, $SD = .59$). Therefore, H1b's assumption that participants in the probabilistic group would report greater perceived severity was not supported by the available data. There was also a significant main effect of exposure. Perceived severity decreased across exposures, with the highest perceptions at Exposure 1 ($M = 3.91$, $SD = .96$) and the lowest at Exposure 3 ($M = 2.59$, $SD = 1.08$). A significant interaction between warning format and exposure was also observed. Follow-up estimated marginal means indicated that probabilistic warnings led to higher severity perceptions at Exposure 1 ($M = 4.32$, $SD = 0.72$) compared to traditional warnings ($M = 3.52$, $SD = 1.0$). However, by Exposure 3, this pattern reversed, with traditional warnings generating higher severity perceptions ($M = 2.98$, $SD = 1.1$) than probabilistic warnings ($M = 2.18$, $SD = 0.89$). Pairwise comparisons with Bonferroni adjustment confirmed significant differences across all exposures ($p < .001$), with perceived severity decreasing across each subsequent exposure in both conditions.

Age moderated the effect of warning format, such that perceptions of severity differed slightly across age groups depending on the warning format viewed. Gender, while not moderating any interaction effects, was a significant overall predictor of perceived severity. Residence status, prior tornado experience, warning frequency, and warning familiarity also emerged as significant main effects, indicating differences in severity perceptions across these characteristics.

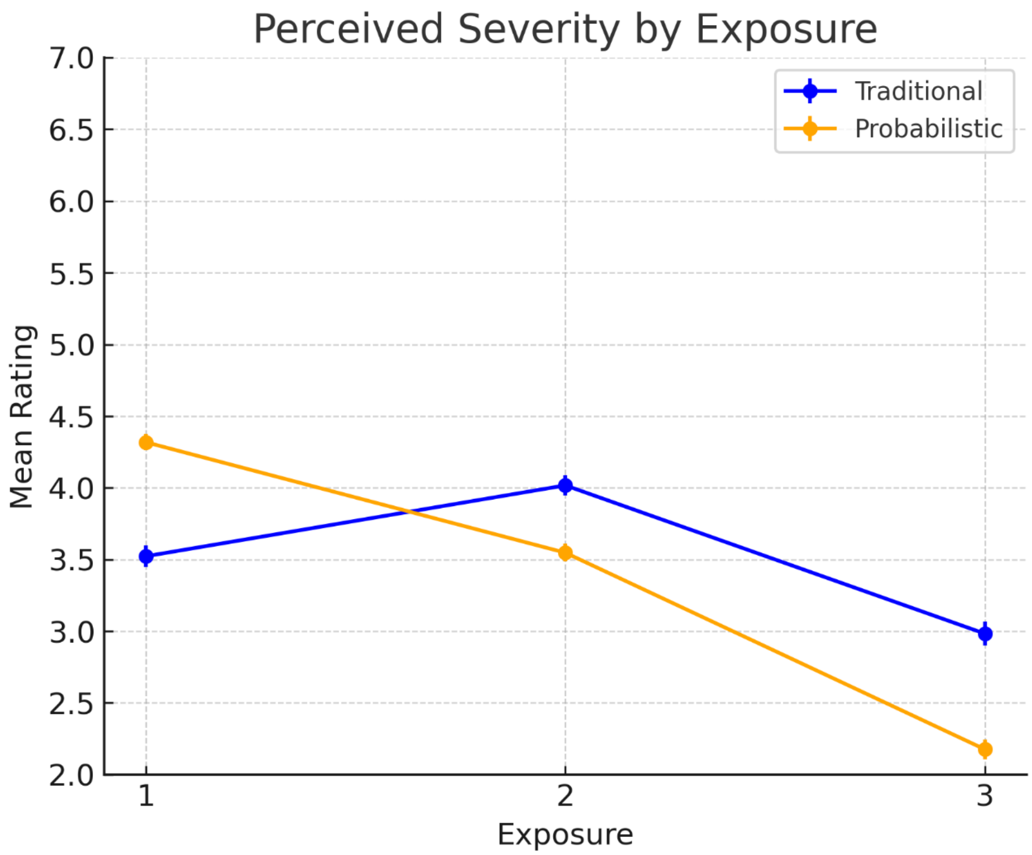

Figure 3: Perceived severity across warning format by exposure

Message Clarity

Results revealed that the main effect of warning format on message clarity was not statistically significant. There was no significant difference in perceived message clarity between the traditional warning format (M = 5.52, SD = 1.33) and the probabilistic warning format (M = 5.24, SD = 1.47). However, there was a significant main effect of exposure. Perceived message clarity decreased across exposures, from Exposure 1 (M = 5.73, SD = 1.47) to Exposure 3 (M = 5.05, SD = 1.65). The interaction between warning format and exposure was not statistically significant, suggesting that the effect of exposure on clarity did not differ between warning formats. Pairwise comparisons with Bonferroni adjustment indicated significant differences between all three exposures ($p < .001$), with clarity decreasing steadily across exposures for both warning formats.

Age showed a significant main effect, indicating that perceptions of message clarity varied slightly across age groups. Residence status also showed a significant main effect, such that participants' living situations were associated with differences in perceived clarity. Additionally, having children moderated the effect of warning format, indicating differing perceptions of message clarity between parents and non-parents depending on the warning format viewed.

Information Sufficiency

Results indicated that the main effect of warning format on information sufficiency was not statistically significant. Participants in the traditional warning condition ($M = 5.44$, $SD = 1.26$) and the probabilistic warning condition ($M = 5.45$, $SD = 1.17$) reported similar perceptions of information sufficiency. Therefore, H1c's assumption that participants in the probabilistic group would report greater perceived information sufficiency was not supported by the available data. There was a significant main effect of exposure. Perceived information sufficiency decreased over time, from Exposure 1 ($M = 5.61$, $SD = 1.35$) to Exposure 3 ($M = 5.25$, $SD = 1.43$). Additionally, there was a significant interaction between warning format and exposure. Follow-up estimated marginal means indicated that probabilistic warnings led to higher perceived information sufficiency at Exposure 1 ($M = 5.76$, $SD = 1.27$) than traditional warnings ($M = 5.46$, $SD = 1.41$). However, this advantage diminished across subsequent exposures, with nearly identical ratings by Exposure 3 (Probabilistic: $M = 5.14$, $SD = 1.39$; Traditional: $M = 5.35$, $SD = 1.46$). Pairwise comparisons with Bonferroni adjustment indicated significant decreases in perceived information sufficiency between Exposure 1 and Exposure 3 ($p < .001$), while the difference between Exposure 1 and 2 was not significant ($p = .089$).

Education moderated the effect of exposure, such that perceptions of information sufficiency across repeated warnings varied slightly by education level. Additionally, having children moderated the effect of warning format by exposure, and warning frequency moderated the effect of warning format alone, indicating differing perceptions based on these characteristics.

Information Seeking

Results showed that the main effect of warning format on information seeking was not statistically significant. Perceived information seeking was nearly identical between the traditional warning condition ($M = 4.87$, $SD = 1.63$) and the probabilistic warning condition ($M = 4.91$, $SD = 1.33$). Therefore, H1d's assumption that participants in the probabilistic group would report lower intentions to seek additional sources of information was not supported by the available data. There was a significant main effect of exposure. Perceived information seeking increased from Exposure 1 ($M = 4.69$, $SD = 1.76$) to Exposure 2 ($M = 5.11$, $SD = 1.68$) and remained relatively stable at Exposure 3 ($M = 4.87$, $SD = 1.68$). The interaction between warning format and exposure was not significant, indicating that the pattern of information seeking across exposures did not differ between the two warning formats. Pairwise comparisons with Bonferroni adjustment indicated a significant increase in information seeking between Exposure 1 and Exposure 2 ($p < .001$), and a significant decrease between Exposure 2 and Exposure 3 ($p = .002$). No significant difference was observed between Exposure 1 and Exposure 3 ($p = .188$).

Gender moderated the combined effects of warning format and exposure, such that patterns of information seeking varied slightly between men and women across repeated warnings. Residence status moderated the combined effects of warning format and exposure, with differences emerging based on participants' living situations. Additionally, having children moderated the effect of warning format alone, indicating variation in information seeking between parents and non-parents depending on the warning format viewed.

Protective Action

Results revealed a significant main effect of exposure, indicating that participants' intention to take protective action changed across warning exposures. Follow-up pairwise comparisons (with Bonferroni adjustment) showed that immediate action scores significantly decreased from Exposure 1 ($M = 4.30$, $SD = .89$) to Exposure 2 ($M = 4.26$, $SD = .90$) to Exposure 3 ($M = 3.30$, $SD = 1.22$). There was also a significant main effect of warning format, with participants exposed to traditional warnings ($M = 4.13$, $SD = .83$) reporting higher immediate action intentions than those exposed to probabilistic warnings ($M = 3.77$, $SD = .80$). Therefore, H1e's assumption that participants in the probabilistic group would report greater intentions to take immediate protective action was not supported by the available data. Additionally, the interaction between warning format and exposure was significant, suggesting that the decline in immediate action intentions over repeated exposures was steeper in the probabilistic warning condition than in the traditional warning condition.

Gender moderated the combined effects of warning format and exposure, such that patterns of immediate action intention varied slightly between men and women across repeated warnings. Residence status moderated the effect of exposure alone, with differences emerging based on participants' living situations. Education moderated the effect of exposure alone, suggesting slight variation in immediate action intentions across warning exposures depending on education level. Prior tornado experience moderated the effect of exposure alone, indicating different patterns between participants with and without prior tornado exposure. Warning familiarity moderated the combined effects of warning format and exposure, such that perceptions of immediate action intention differed across levels of familiarity with warning messages.

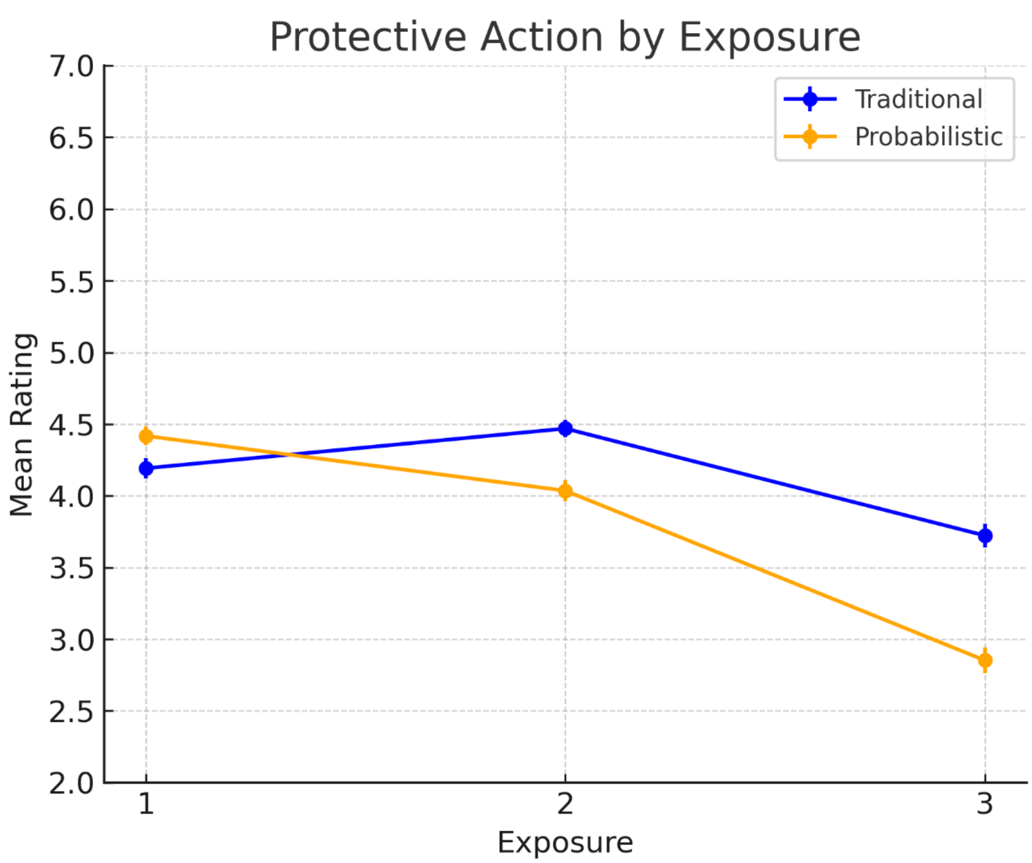

Figure 4: Protective action intentions across warning format by exposure

Cognitive Load

A mixed-effects repeated-measures model was conducted to examine whether cognitive load varied as a function of warning format (traditional vs. probabilistic), exposure order (first, second, third warning), and their interaction. Results indicated a significant main effect of exposure. Cognitive load was highest at Exposure 1 ($M = 4.77$, $SD = 2.06$) and decreased slightly by Exposure 2 ($M = 4.68$, $SD = 2.16$) and Exposure 3 ($M = 4.36$, $SD = 1.98$). However, there was no significant main effect of warning format. Therefore, H2a's assumption that participants in the probabilistic group would report higher cognitive load was not supported by the available data. Additionally, the interaction between warning format and exposure was not significant.

Age moderated the effect of warning format, such that perceptions of cognitive load differed slightly across age groups depending on the warning format viewed. Education moderated the effect of exposure alone, and having children moderated the effect of exposure alone, suggesting that perceptions of cognitive load across repeated warnings varied depending on participants' education level and parental status.

H2b surmised that, regardless of group, participants who reported higher cognitive load would also report lower likelihood of taking immediate protective action. A Pearson correlation was conducted to explore this relationship. Results indicated a non-significant positive correlation between cognitive load and protective action. Therefore, H2b was not supported by the available data.

Perceived Message Sensation Value

To assess participants' perceptions of the warning design, three distinct constructs were measured following exposure to all three warnings: emotional arousal, dramatic impact, and novelty. Each construct was measured using items averaged across the three warning exposures (e.g. by averaging all three emotional arousal responses within the standard warning group). This resulted in composite scores for emotional arousal, dramatic impact, and novelty. A factor analysis was conducted to determine whether these dimensions could be combined into a single PMSV composite. However, results indicated poor communalities for novelty (.087) and dramatic impact (.364), leaving only emotional arousal with a sufficiently high communality score (.827). Therefore, the emotional arousal, dramatic impact, and novelty were analyzed as separate dependent variables rather than combined into a single PMSV score.

Next, a MANOVA was conducted to examine whether the format of the warning influenced participants' perceptions of emotional arousal, dramatic impact, and novelty. Given the violation of the assumption of homogeneity of covariance matrices (*B*ox's $M = 24.55$, $p < .001$), Pillai's Trace was utilized. Results indicated a significant multivariate effect of the warning's format on PMSV variables.

Follow-up univariate tests indicated that participants exposed to probabilistic warnings reported significantly greater emotional arousal, dramatic impact, and novelty than those exposed to traditional warnings (Fig. 5), providing consistent support for H3a, H3b, and H3c. A summary of detailed statistics is provided in Table 2. Specifically, participants exposed to probabilistic warnings reported greater arousal ($M = 5.63$, $SD = 1.16$) than those who viewed traditional warnings ($M = 5.01$, $SD = 1.54$), greater dramatic impact ($M = 5.41$, $SD = 1.36$) compared to traditional warnings ($M = 4.59$, $SD = 1.77$) and greater novelty ($M = 4.17$, $SD = 1.64$) compared to traditional warnings ($M = 3.77$, $SD = 1.56$).

Dimension	Trad. M	Prob. M	$F(1{,}344)$	p	η_p^2
Emotional Arousal	5.01 [4.81,5.21]	5.63 [5.43,5.84]	17.93	<0.001	.050
Dramatic Impact	4.59 [4.35,4.82]	5.41 [5.17,5.65]	23.43	<0.001	.063
Novelty	3.77 [3.53,4.01]	4.17 [3.93,4.41]	5.49	.020	.016

Table 2: Univariate ANOVA results for PMSV variables

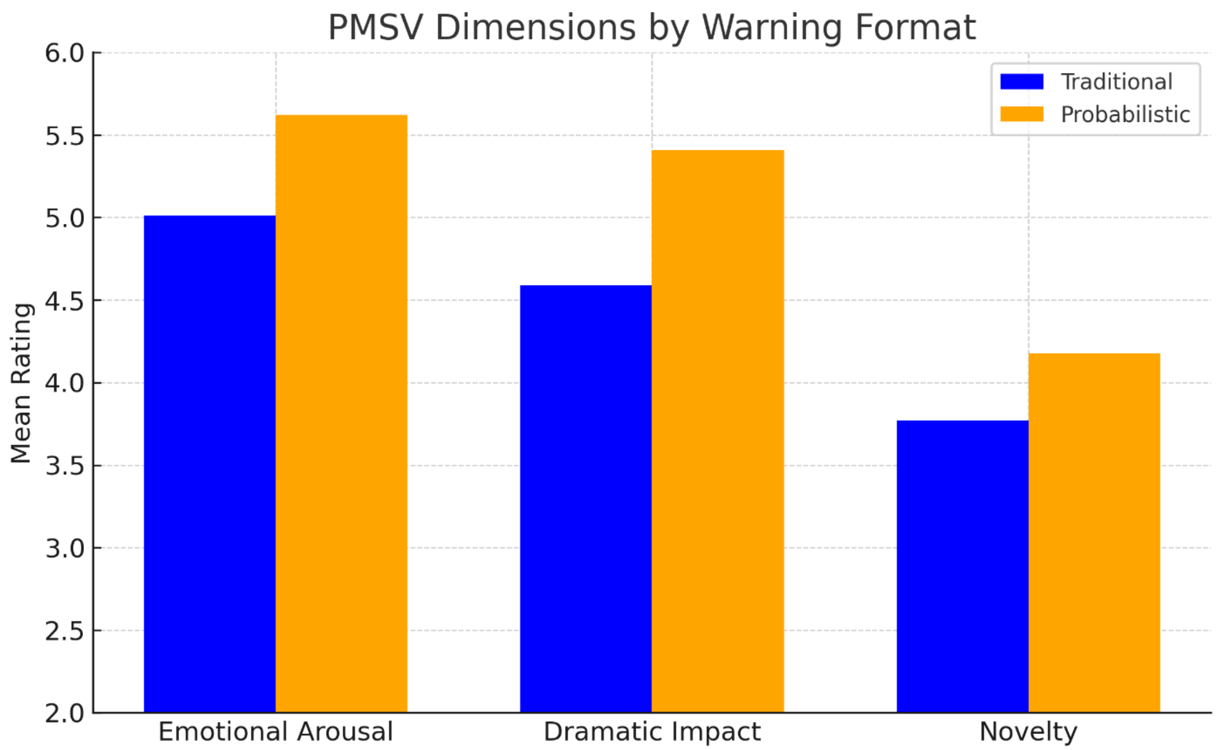

Figure 5: PMSV dimensions across warning format

H3d hypothesized that participants exposed to probabilistic warnings would rate them as more accurate compared to those who viewed traditional warnings. Accuracy was measured by a single reverse-coded item assessing whether the warning display made participants question its accuracy. An independent samples t-test was conducted to compare the groups. Results indicated that there was not a statistically significant difference in perceived accuracy between the traditional warning group (*M* = 4.86, *SD* = 1.62) and the probabilistic warning group (*M* = 4.99, *SD* = 1.47). Therefore, H3d was not supported by the available data.

H3e hypothesized that participants exposed to probabilistic warnings would report a greater sense of urgency to act after all three exposures compared to those who viewed traditional warnings. This behavioral intention was measured using a reverse-coded question stating that the visual display of the warning made them feel less urgency to act. An independent samples t-test was conducted to compare the groups. Results indicated a statistically significant difference in behavioral intention. Specifically, participants in the probabilistic warning group (*M* = 5.34, *SD* = 1.43) reported a significantly higher intention to act compared to those in the traditional warning group (*M* = 4.93, *SD* = 1.61). Therefore, H3e was supported by the available data.

Message Fatigue

H4a posited that participants with greater message fatigue would report lower trust in forecasters due to repeated false alarms. Trust was assessed with a single item measuring reduced trust in forecasters due to repeated false alarms. A Pearson correlation was conducted to analyze the data. Counterintuitively, results indicated a statistically significant positive relationship between message fatigue and trust. Since H4a suggested a negative relationship, it was not supported by the available data.

To test H4b, a hierarchical regression analysis was conducted to examine whether message fatigue moderated the effect of warning format on perceived likelihood. Message fatigue was mean-centered, and an interaction term was created by multiplying centered message fatigue by warning format. In the first step, centered message fatigue and warning format were entered as predictors. The resultant model significantly predicted likelihood perceptions, R^2 = .055, $F(2, 343) = 9.96$, $p < .001$. Both message fatigue ($B = -0.174$, $p < .001$) and warning format ($B = -0.214$, $p = .031$) were significant predictors. Higher fatigue was associated with lower likelihood perceptions, and participants in the probabilistic group reported lower likelihood perceptions than those in the traditional group. In the second step, the interaction term was added to the model. However, this addition did not significantly improve the model fit, $\Delta R^2 = .001$, $F(1, 342) = 0.51$, $p = .474$, and the interaction effect alone was not statistically significant, $B = -.066$, $p = .474$. In the third step, a bootstrapping procedure with 5,000 samples was conducted to assess the robustness of the interaction effect given the low variance in message fatigue. The bootstrapped 95% confidence interval for the interaction term ranged from -0.242 to .099, which includes 0, indicating the interaction effect was not robust.

Taken together, these results did not support H4b. Although both message fatigue and warning format independently predicted perceived likelihood, the interaction between them was not significant. This suggests that message fatigue did not moderate the relationship between warning format and perceived likelihood of a tornado.

To test H4c, a hierarchical regression analysis was conducted to examine whether message fatigue moderated the effect of warning format on emotional arousal, dramatic impact, and novelty. Message fatigue was mean-centered to create a centered message fatigue score, and an interaction term was created by multiplying centered message fatigue by warning format. The robustness of the interaction effect was assessed through a bootstrapping procedure with 5,000 samples. The 95% confidence interval in all cases included 0, indicating the interaction effort was not significant. Detailed statistics are referenced in Table 3.

Outcome	Fatigue B (p)	Format B (p)	Interaction B (p)	R^2	ΔR^2
Emotional arousal	-0.414 (<.001)	0.663 (<.001)	0.134 (0.306)	.148	.003
Dramatic impact	-0.513 (<.001)	0.883 (<.001)	0.023 (0.877)	.175	.000
Novelty	0.057 (0.481)	0.401 (0.020)	0.215 (0.184)	.018	.005

Table 3: Format × Fatigue effects on PMSV

The resultant model significantly predicted emotional arousal, dramatic impact, and novelty. Message fatigue and warning format were significant predictors for both emotional arousal and dramatic impact. For novelty, the warning format was a significant predictor, but message fatigue was not. In all cases, adding the interaction term to the model did not significantly improve model fit, and the interaction effects were not statistically significant. Together, these findings did not support H4c. While message fatigue and warning format each played independent roles in shaping participants' perceptions of PMSV outcomes, message fatigue did not moderate the effect of warning format on emotional arousal, dramatic impact, or novelty.

Desensitization

H4d hypothesized that participants exposed to probabilistic warnings would report lower levels of desensitization compared to those exposed to traditional warnings. An independent samples t-test was conducted to analyze the data. Results indicated no significant difference in desensitization between the traditional group (M = 2.45, SD = 1.28) and the probabilistic group (M = 2.56, SD = 1.46), $t(345)$ = -0.76, p = .450, d = -0.08, 95% CI [-0.29, 0.13]. Therefore, H4d was not supported by the available data.

Discussion

This study examined traditional versus probabilistic tornado warning formats across a range of PADM constructs—namely, perceived likelihood, severity, message clarity, information sufficiency and seeking, and protective action intention—as well as message fatigue, cognitive load, PMSV, and desensitization. A 2x3 mixed design was utilized within a Qualtrics survey of 347 participants to explore these variables. Broadly, the findings indicate that probabilistic warnings generated higher initial perceptions of risk and emotional response but declined more sharply over repeated exposures compared to traditional warnings.

Results from H1a and H1b indicated that traditional warnings ultimately produced higher perceptions of likelihood and severity after multiple exposures, contradicting the expectation that probabilistic warnings would excel in these metrics. Perceived message clarity and information sufficiency (H1c), meanwhile, eroded over time regardless of warning format, though probabilistic warnings held a temporary advantage in initial exposure. There were no significant differences in information seeking (H1d), though exposure effects dominated, with participants most likely to seek information at Exposure 2 before declining at Exposure 3. Again, contrary to expectations, traditional warnings generated stronger and more stable responses to protective action intention (H1e), while probabilistic warnings suffered a sharp decline over time.

H2a hypothesized that probabilistic tornado warnings would elicit greater cognitive load, while H2b hypothesized that as cognitive load increased, protective action intention would decline. However, neither hypothesis was supported by the available data. As it related to the core principles of PMSV, H3a through H3c were supported, as probabilistic warnings elicited greater emotional arousal, dramatic impact, and novelty. While the probabilistic warnings were not perceived as more accurate in a rebuke of H3d, they did increase the perceived urgency to act in support of H3e. Importantly, H4a was an unexpected finding that found higher levels of message fatigue among participants was associated with greater trust in forecasters. H4b and H4c found that message fatigue predicted lower perceptions of likelihood and PMSV outcomes but did not significantly moderate the warning format effects. H4d was not supported, as desensitization did not differ between formats Across the core PADM variables, demographic variables only infrequently moderated the effects of warning format or exposure. These effects were inconsistent across models, though some characteristics such as education, parental status, residence status, prior tornado experience, and warning frequency appeared more frequently than others.

The PADM framework posits that threat perceptions are fundamental in driving protective action decisions (Lindell & Perry, 2012). In my study, I operationalized threat perception through measures of perceived likelihood and severity. The results showed that traditional warnings, compared to probabilistic warnings, generated higher ratings for both likelihood and severity. Consequently, the PADM would predict that traditional warnings should produce stronger protective action intentions, a prediction that my data supports. Additionally, the PADM emphasizes the role of information sufficiency in influencing decision-making. The initial advantage of probabilistic warnings in conveying sufficient information suggests they initially improved comprehension; however, this advantage faded over repeated exposures, undermining sustained protective responses. These outcomes underscores the PADM's assertion that multiple, interrelated components must align and be sustained to drive effective protective action.

PMSV refers to the degree to which warning attributes evoke sensory, affective, and arousal responses in individuals (Everett & Palmgreen, 1995). Prior studies indicate that messages high in PMSV capture attention and elicit strong emotional responses. In my study, we operationalized PMSV using three dimensions: emotional arousal, dramatic impact, and novelty. The results show that probabilistic warnings are significantly more effective than traditional warnings at generating these responses. Although heightened PMSV was not accompanied by an increased immediate intention to act (H1e), the data did reveal that participants felt a stronger general urge to take protective action (H3e). This pattern is consistent with the PADM's focus on warning characteristics: while a more emotionally charged message can elevate concern about the threat, it does not automatically ensure a corresponding behavioral response. In other words, probabilistic warnings may prime individuals with a heightened sense of urgency and concern, even if this does not immediately translate into action. Such findings underscore the notion that multiple factors must interact and be sustained to drive protective behavior.

Cognitive load refers to the mental effort required to process and comprehend information (Paas, 1992). Cognitive load theory suggests that when information is complex or unfamiliar, the high mental demand can hinder comprehension and subsequent decision-making (Sweller, 1998). In my study, cognitive load was assessed using a single-item measure and served as a proxy for the effort participants had to invest in interpreting the warning messages. My results indicate that cognitive load was highest at the initial exposure and then decreased slightly with subsequent exposures. Notably, there was no significant difference in cognitive load between traditional and probabilistic warning formats. This is an encouraging result that suggests communicating threat probabilistically does not overwhelm participants and stifle decision-making, a concern that has plagued academics for decades (Joslyn & LeClerc, 2012).

Message fatigue is generally defined as the weariness and reduced motivation to process warning messages following repeated exposure (So et al., 2017). In theory, high message fatigue is expected to erode trust in the warning source and lower protective action intentions. In my study, however, the results for message fatigue deviated from these expectations. Specifically, although we hypothesized that higher fatigue would be associated with lower trust in forecasters (H4a), my data revealed a significant positive relationship. This unexpected result implies that elevated message fatigue may not signal diminished trust in forecasters; rather, it may reflect an increased familiarity with severe weather that, in turn, fosters greater reliance and trust.

Additionally, hierarchical regression analyses aimed at testing whether message fatigue moderated the relationship between warning format and perceived likelihood or PMSV outcomes (H4b-H4c) found no significant interaction effects. This indicates that regardless of the level of fatigue, the differences between traditional and probabilistic warnings remained consistent, and its moderating influence on key pre-decisional processes appears minimal in my study. Relatedly, desensitization, or the diminished emotional response to warnings after repeated exposure (Cho & Salmon, 2007), was measured as part of my study but did not show significant differences across warning formats. The straightforward explanation for this result is that desensitization is considered a byproduct of high message fatigue (So et al., 2017); because fatigue levels were relatively low in this dataset, there was no opportunity for the effects of desensitization to materialize. This is important because it challenges a widespread perception among meteorologists that over-warning may lead to disengagement from the public.

In practice, these results indicate the crucial role that every aspect of the warning plays in driving behavior change. It is not sufficient for a warning to simply be dramatic, fresh, or emotional. This must be paired with clear messaging that provides sufficient information and a clear picture of threat perception to spur citizens to act. These findings suggest that one promising approach may be to develop a hybrid warning design that combines the simplicity and clarity of traditional warnings with probabilistic elements—applied when perceived threat is high—to maintain public attention. Further, these findings challenge the notion that forecasters should issue tornado warnings sparingly for fear of fatiguing the public. While forecasters should always strive for accurate warnings with few false alarms, the public seems broadly supportive of their mission to keep citizens safe, even if the warnings are not always perfect.

While my study provided valuable insights into the temporal dynamics of tornado warning perceptions using a repeated-measures design, several limitations must be acknowledged. Although we strived to recruit a more diverse sample (n = 347) covering ages 18 to 75 (M = 32.9), my sample was still disproportionately White (75%) and college-educated (85.7% with some college education or higher). This raises important questions about the generalizability of these findings to minority and non-college-educated groups where differences in cultural background or educational attainment may influence how the warnings are processed.

The reliance on self-reported Likert scales also introduces inherent bias. These measures captured participants' perceptions in a hypothetical, online setting rather than in a genuine life-threatening scenario. Therefore, there is a risk that the responses may not fully reflect how individuals would behave in an actual tornado emergency. Future research should explore alternative methodologies that introduce real stakes, such as laboratory simulations. These simulations should also consider employing behavioral or physiological measures such as eye tracking, heart rate, and respiration for a more nuanced understanding of instantaneous reactions. In addition, while useful for directly observing how warning perceptions evolve over time, my repeated-exposure design was constrained. Participants were only exposed to warnings at three fixed locations within the warning area, which may not have allowed for a comprehensive picture of how participants process these types of warnings. Future studies might utilize more dynamic or interactive exposure methods.

Finally, there are limitations regarding the measurement of message fatigue and cognitive load. The message fatigue scale used in this study was originally developed for health communication contexts, where messages are encountered over much longer periods and at higher frequencies. In the context of tornado warnings, which are comparatively rare, the scale may not fully capture acute fatigue responses. Future research should explore whether a new message fatigue scale should be adopted to cover shorter temporal windows such as explored here. Similarly, we employed a single-item measure to assess cognitive load because this remains the most common way to measure the concept (Sweller, 2018). However, newer multi-item scales may nevertheless offer a more nuanced understanding of the cognitive effort required to process warning messages, and future research should consider employing these scales.

Conclusion

In sum, probabilistic tornado warnings reliably increase emotional arousal, dramatic impact, and novelty without adding cognitive load, and they momentarily boost perceived risk at the first exposure. However, by the third exposure that risk advantage vanishes, and traditional warnings again drive the highest threat perceptions and strongest protective action intentions. Unexpectedly, greater message fatigue was linked to increased trust in forecasters rather than disengagement. By integrating message fatigue and PMSV into a meteorological context alongside the PADM, this work refines theoretical understanding and points toward hybrid warning designs that blend the clarity of traditional formats with the dynamic specificity of probabilistic elements.

However, these findings rest on a predominantly White, college-educated sample, self-reported Likert measures, fixed simulated exposures, and single-item scales for cognitive load and fatigue, all of which are constraints that may limit generalizability. Future research should diversify participant pools, employ laboratory or field simulations with behavioral and physiological metrics (e.g., eye-tracking, heart rate), and adopt multi-item scales to more fully capture cognitive effort and message fatigue. Extending this integrative approach with additional risk communication and cognitive-psychology frameworks will further guide the development of effective tornado warnings—and ultimately enhance public safety.

References

Ahlstrom, U. (2005). Work domain analysis for air traffic controller weather displays. *Journal of Safety Research, 36*(2), 159-169. https://doi.org/10.1016/j.jsr.2005.03.001

Ahlstrom, U., & Dworsky, M. (2012). *Effects of weather presentation symbology on general aviation pilot behavior, workload and visual scanning* (*DOT/FAA/TC-12/55)*. Federal Aviation Administration. https://rosap.ntl.bts.gov/view/dot/71733

Ahsan, M., Takeuchi, K., Vink, K., Ohara, M. (2016). A systematic review of the factors affecting the cyclone evacuation decision process in Bangladesh. *Journal of Disaster Research, 11*(4), 742-753. http://dx.doi.org/10.20965/jdr.2016.p0742

Annetta, L., & Newton, M. (2023). Perceived cognitive load of extended reality serious educational games about climate change. *Emerging Trends in Education, 7*, 25-43. Retrieved from https://docensjournal.org/index.php/docens/article/view/55

Ash, K., Schuman, R., & Bowser, G. (2014). Tornado warning trade-offs: Evaluating choices for visually communicating risk. *Weather, Climate, and Society, 6*(1), 104-118. https://doi.org/10.1175/WCAS-D-13-00021.1

Brooks, H., & Correia, J. (2018). Long-term performance metrics for National Weather Service tornado warnings. *Weather and Forecasting, 33*(6), 1501-1511. https://doi.org/10.1175/WAF-D-18-0120.1

Calvo, L., Christel, I., Terrado, M., Cucchietti, F., & Pérez-Montoro, M. (2022). Users' cognitive load: A key aspect to successfully communicate visual climate information. *Bulletin of the American Meteorological Society, 103*(1), E1-E16. https://doi.org/10.1175/BAMS-D-20-0166.1

Cho, H., & Salmon, C. (2007). Unintended effects of health communication campaigns. *Journal of Communication, 57*(2), 293-317. http://dx.doi.org/10.1111/j.1460-2466.2007.00344.x

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research, 169*, 323-338. https://doi.org/10.1016/S0079-6123(07)00020-9

Donohew, L., Lorch, E., & Palmgreen, P. (1998). Applications of a theoretic model of information exposure to health interventions. *Human Communication Research, 24*(3), 454-468. https://doi.org/10.1111/j.1468-2958.1998.tb00425.x

Downs, A. (1972). *Up and down with ecology—The issue-attention tragedies*. Public Interest, 28, 28-50.

Erdman, J. (2021, April 6). *How accurate are tornado warnings? Here's what the last five years of data says*. The Weather Channel. https://weather.com/storms/tornado/news/2021-04-05-tornado-warning-nws-accuracy

Everett, M., & Palmgreen, P. (1995). Influences of sensation seeking, message sensation value, and program context on effectiveness of anticocaine public service announcements. *Health Communication*, 7(3), 225-248. https://doi.org/10.1207/s15327027hc0703_3

Firth, J., Torous, J. Stubbs, B., Firth, J. Steiner, G., Smith, L., Alvarez-Jimenz, M., Gleeson, J., Vancampfort, D., Armitage, C., & Sarris, J. (2019). The "online brain": How the internet may be changing our cognition. *World Psychiatry, 18*(2), 119-129. https://doi.org/10.1002/wps.20617

Fox, P., Edwards, J., Snyder, F., Makinson, K., & Hamby, D. (2013). The effect of cognitive load on decision making with graphically displayed uncertainty information. *Risk Analysis, 34*(8), 1495-1505. http://dx.doi.org/10.1111/risa.12161

Hanson-Easey, S., Hansen, A., Williams, S., & Bi, P. (2019). *Communicating about heatwaves: risk perception, message fatigue, and threat normalisation*. School of Public Health, Faculty of Health and Medical Sciences, The University of Adelaide.

Harrington, N., Lane, D., Donohew, L., Zimmerman, R., Holmes, G., An, J., Cheah, W., McClure, L., Buckingham, T., Garofalo, E., & Bevins, C. (2003). Persuasive strategies for effective anti-drug messages. *Communication Monographs, 70*(1), 16-38. http://dx.doi.org/10.1080/03637750302478

Huang, S.H., Lindell, M., & Prater, C. (2016). Multistage model of hurricane evacuation decision: Empirical study of hurricanes Katrina and Rita. *Natural Hazards Review, 18*(3). https://doi.org/10.1061/(ASCE)NH.1527-6996.0000237

Hullman, J (2020). Why authors don't visualize uncertainty. *IEEE Transactions on Visualization and Computer Graphics, 26*(1), 130-139. https://doi.org/10.1109/TVCG.2019.2934287

Islam, K., Sopory, P., & Seeger, M. (2024). Compound crisis communication and household preparedness: Examining the effects of evidence type and crisis message fatigue. *Journal of Contingencies and Crisis Management, 32*(3). https://doi.org/10.1111/1468-5973.12609

Jon, I., Huang, S., & Lindell, M. (2018). Perceptions and reactions to tornado warning polygons: Would a gradient polygon be useful? *International Journal of Disaster Risk Reduction, 30*, 132-144. https://doi.org/10.1016/j.ijdrr.2018.01.035

Jon, I., Huang, S., & Lindell, M. (2019). Perceptions and expected immediate reactions to severe storm displays. *Risk Analysis, 39*(1), 274-290. https://doi.org/10.1111/risa.12896

Jonides, J., Lacey, S., & Nee, D. (2005). Processes of working memory in mind and brain. *Current Directions in Psychological Science, 14*(1), 2-5. https://doi.org/10.1111/j.0963-7214.2005.00323.x

Joslyn, S., & LeClerc, J. (2012). Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. *Journal of Experimental Psychology: Applied, 18*(1), 126-140. http://dx.doi.org/10.1037/a0025185

Keating, D., Skurka, C. (2024). Meta-analytic evidence that message fatigue is associated with unintended persuasive outcomes. *Communication Research*. https://doi.org/10.1177/00936502241287875

Khamaj, A., Kang, Z., & Argyle, E. (2019). Users' perceptions of smartphone weather applications' usability. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 63*(1), 2216-2220. https://doi.org/10.1177/1071181319631098

Kinnick, K., Krugman, D., & Cameron, G. (1996). Compassion fatigue: Communication and burnout toward social problems. *Journal & Mass Communication Quarterly, 73*(3), 687-707. http://dx.doi.org/10.1177/107769909607300314

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.01997

Klockow, K. (2013). *Spatializing tornado warning lead-time: Risk perception and response in a spatio-temporal framework* [Doctoral dissertation, University of Oklahoma]. PQDT Open. https://www.proquest.com/openview/f902aeedf05f998cb0fa953de3afa208/1

KY3 Staff. (2022, May 21). *Eleven years later: Remembering the May 22, 2011 tornado in Joplin*. KY3. https://www.ky3.com/2022/05/21/eleven-years-later-remembering-may-22-2011-tornado-joplin/

Lang, A., Newhagen, J., & Reeves, B. (1996). Negative video as structure: Emotion, attention, capacity, and memory. *Journal of Broadcasting & Electronic Media, 40*(4), 460-477. https://doi.org/10.1080/08838159609364369

Lazarsfeld, P., & Merton, R. (1948). *Mass communication, popular taste, and organised social action*. In L. Bryson (Ed.), The communication of ideas (pp. 95–118). New York: Harper & Row.

Lim, J., Liu, B., & Egnoto, M. (2019). Cry wolf effect? Evaluating the impact of false alarms on public responses to tornado alerts in the Southeastern United States. *Weather, Climate, and Society, 11*(3), 549-563. https://doi.org/10.1175/WCAS-D-18-0080.1

Lin, S. (2013). Perceived impact of a documentary film: An investigation of the first-person effect and its implications for environmental issues. *Science Communication, 35*(6), 708-733. https://doi.org/10.1177/1075547013478204

Lindell, M. (2020). Improving hazard map comprehension for protective action decision making. *Frontiers of Computer Science, 2*. https://doi.org/10.3389/fcomp.2020.00027

Lindell, M., & Perry, R. (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

Liu, R., Wang, L., Koszalka, T., & Wan, K. (2022). Effects of immersive virtual reality classrooms on students' academic achievement, motivation and cognitive load in science lessons. *Journal of Computer Assisted Learning, 38*(5), 1422-1433. http://dx.doi.org/10.1111/jcal.12688

Mackie, B. (2014). *Warning fatigue: Insights from the Australian bushfire context* [Thesis, University of Canterbury]. University of Canterbury Library. https://ir.canterbury.ac.nz/items/46203cf9-3397-470a-b63b-122e0f1c3b44

Miran, S., Ling, C., Rothfusz., L. (2017). Effective method to convey threat information for tornado: Probabilistic hazard information vs deterministic hazard information. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*(1), 292-296. https://doi.org/10.1177/1541931213601554

Miran, S., Ling, C., Gerard, A., & Rothfusz, L. (2018). The effect of providing probabilistic information about a tornado threat on people's protective actions. *Natural Hazards, 94*, 743-758. https://doi.org/10.1007/s11069-018-3418-5

Miran, S. M., Ling, C., Gerard, A., & Rothfusz, L. (2019). Effect of providing the uncertainty information about a tornado occurrence on the weather recipients' cognition and protective action: Probabilistic hazard information versus deterministic warnings. *Risk Analysis, 39*(7), 1533-1545. https://doi.org/10.1111/risa.13289

Miran, S., Ling, C., & James, J. (2020). People's thresholds of decision-making against a tornado threat using dynamic probabilistic hazard information. *International Journal of Disaster Risk Reduction, 42*. https://doi.org/10.1016/j.ijdrr.2019.101345

Morgan, S., Palmgreen, P., Stephenson, M., Hoyle., R., & Lorch, E. (2003). Associations between message features and subjective evaluations of the sensation value of antidrug public service announcements. *Journal of Communication, 53*(3), https://doi.org/10.1111/j.1460-2466.2003.tb02605.x

National Weather Service. (n.d.). *Understand tornado alerts*. https://www.weather.gov/safety/tornado-ww

Noar, S., Palmgreen, P., Zimmerman, R., Lustria, M., & Lu, H. (2010). Assessing the relationship between perceived message sensation value and perceived message effectiveness: Analysis of PSAs from an effective campaign. *Communication Studies, 61*(1), 21-45. https://doi.org/10.1080/10510970903396477

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429-434. https://doi.org/10.1037/0022-0663.84.4.429

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Palmgreen, P., & Donohew, L. (2003). *Effective mass media strategies for drug abuse prevention campaigns*. Springer. http://dx.doi.org/10.1007/0-387-35408-5_2

Paul, B., Stimers, M., & Caldas, M. (2015). Predictors of compliance with tornado warnings issued in Joplin, Missouri, in 2011. *Disasters, 39*(1), 108-124. https://doi.org/10.1111/disa.12087

Qin, C., Joslyn, S., Savelli, S., Demuth, J., Morss, R., & Ash, K. (2024). The impact of probabilistic tornado warnings on risk perceptions and responses. *Journal of Experimental Psychology: Applied, 30*(1), 206-239. https://doi.org/10.1037/xap0000486

Schmid, R. (2011, May 30). *Despite advances in warning systems, tornadoes are still a deadly threat*. Tuscaloosa News. https://www.tuscaloosanews.com/story/news/2011/05/30/despite-advances-in-warning-systems-tornadoes-are-still-a-deadly-threat/28379116007/

Sherman-Morris, K. (2005). Tornadoes, television, and trust—A closer look at the influence of the local weathercaster during severe weather. *Global Environment Change Part B: Environmental Hazards, 6*(4), 201-210. https://doi.org/10.1016/j.hazards.2006.10.002

So, J., Kim, S., & Cohen, H. (2017). Message fatigue: Conceptual definition, operationalization, and correlates. *Communication Monographs, 84*(1), 5-29, http://dx.doi.org/10.1080/03637751.2016.1250429

Stephenson, M., & Palmgreen, P. (2001). Sensation seeking, perceived message sensation value, personal involvement, and processing of anti-marijuana PSAs. *Communication Monographs, 68*(1), 49-71. https://doi.org/10.1080/03637750128051

Stumpf, G., & Gerard, A. (2021). National Weather Service severe weather warnings as threats-in-motion. *Weather and Forecasting, 36*(2), 627-643. https://doi.org/10.1175/WAF-D-20-0159.1

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257-285. https://doi.org/10.1016/0364-0213(88)90023-7

Sweller, J. (2018). Measuring cognitive load. *Perspectives on Medical Education, 7*(1), 1-2. https://doi.org/10.1007/s40037-017-0395-4

Sweller, J., Van Merrienboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296. https://doi.org/10.1023/A:1022193728205

Sweller, J., Van Merriënboer, J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*, 261-292. https://doi.org/10.1007/s10648-019-09465-5

Takeuchi, K., Vink, K., Ohara, M. (2016). A systematic review of the factors affecting the cyclone evacuation decision process in Bangladesh. *Journal of Disaster Research, 11*(4), 742-753. http://dx.doi.org/10.20965/jdr.2016.p0742

Thorson, E., & Lang, A. (1992). The effects of television videographics and lecture familiarity on adult cardiac orienting responses and memory. *Communication Research, 19*(3), 346-349. https://doi.org/10.1177/009365092019003003

Williams, C. (2020). Should severe weather graphics wear a uniform? Exploring the effects of visual and spatial inconsistencies on end user risk perception, uncertainty, and behavioral intentions. [Doctoral dissertation, The University of Georgia]. https://openscholar.uga.edu/record/5273?ln=en

Xu, J. (2015). Message sensation and cognition values: Factors of competition or integration? *Health Communication, 30*(6), 589-597. https://doi.org/10.1080/10410236.2013.852460

Zuckerman, M. (1988). *Behavior and biology: Research on sensation seeking and reactions to the media*. Lawrence Erlbaum Associates.