Combining Probabilistic Hazard Information Forecast Graphics with Wireless Emergency Alert Messages: An Exploratory, Qualitative Study

HAMILTON BEAN,[a] KENSUKE TAKENOUCHI,[b] AND ANA MARIA CRUZ[c]

[a] *University of Colorado Denver, Denver, Colorado*
[b] *Kagawa University, Kagawa, Japan*
[c] *Kyoto University, Kyoto, Japan*

(Manuscript received 3 December 2022, in final form 13 July 2023, accepted 1 August 2023)

ABSTRACT: Since 2019, National Weather Service (NWS) offices have been able to issue 360-character Wireless Emergency Alert ("WEA360") messages for tornadoes. NWS is now considering changing from a "deterministic" to a "probabilistic" warning paradigm. That change could possibly influence how WEA360 messages for tornado are issued in the future. Recent experimental studies have found that probabilistic hazard information (PHI) forecast graphics improve consumers' risk perception for tornadoes, but findings from these studies concerning whether PHI forecast graphics improve people's protective action decision-making are mixed. The present study therefore investigated how mock PHI-enhanced WEA360 messages might influence people's risk perception and protective action decision-making. Analysis of qualitative data gathered from a combination of questionnaire and focus group interview methods conducted in collaboration with 31 community members in Denver, Colorado, indicated that inclusion of PHI forecast graphics within WEA360 messages elicited high levels of understanding and message believability but did not consistently lead to appropriate precautionary intent. Because warning response is a complex social phenomenon, PHI may not significantly improve protective action decision-making if PHI forecast graphics are eventually presented to consumers via the Wireless Emergency Alerts system. Factors that PHI stakeholders should consider before the adoption of PHI-enhanced WEA360 messages for consumers are discussed.

SIGNIFICANCE STATEMENT: This study examines how consumers respond to and talk about mock WEA360 messages for tornadoes that contain embedded PHI forecast graphics. As NWS considers moving to a probabilistic warning paradigm, stakeholders will need to determine how PHI forecast graphics might be communicated directly to consumers, if at all. Our findings suggest that combining WEA360 messages with PHI forecast graphics creates challenges and complexities related to consumers' assessment of personal risk and protective action decision-making. Overall, the study suggests that any future PHI-enhanced WEA360 messages provided directly to consumers, if at all, must avoid discrepancies (even subtle) between the level of risk represented by the PHI forecast graphic and the protective action guidance included in the text of the messages.

KEYWORDS: Social science; Probability forecasts/models/distribution; Community; Communications/decision-making; Emergency preparedness

1. Introduction

The U.S. Wireless Emergency Alerts (WEA) system was launched in 2012. Since then, more than 70 000 WEA messages have been issued throughout the United States to warn millions of wireless customers about imminent threats (e.g., floods, wildfires, and hurricanes), abducted children [America's Missing: Broadcast Emergency Response (AMBER) alerts], and other emergencies (e.g., COVID-19). In 2016, the Federal Communications Commission (FCC) increased the number of characters permitted in WEA messages from 90 to 360. The National Weather Service (NWS) subsequently developed "WEA360" message templates for various weather hazards and began issuing longer messages in 2019. These WEA360 messages include additional information about the hazard and recommended protective actions. NWS forecasters can currently issue WEA360 messages for tornadoes. However, older wireless handsets may cause some recipients to receive only a 90-character version of the message. On 30 March 2023, the FCC issued a Further Notice of Proposed Rulemaking (FCC 2023), which indicated that both embedded multimedia and personal location marker could be on the horizon for the WEA system in the future, although both technical and policy changes would be needed before implementation.

In parallel to WEA system developments, the National Oceanic and Atmospheric Administration's (NOAA) National Severe Storms Laboratory (NSSL) has spearheaded the Forecasting a Continuum of Environmental Threats (FACETs) program. FACETs includes an array of subprojects aimed at developing the United States' "next-generation severe weather forecast and warning framework" (National Severe Storms Laboratory 2020, para. 1). A principal focus of FACETs is generating probabilistic hazard information (PHI) forecasting tools and graphics to support NWS's warning activities.

Corresponding author: Hamilton Bean, hamilton.bean@ucdenver.edu

DOI: 10.1175/WCAS-D-22-0140.1

Officials and scholars have contemplated the issuance of PHI forecast graphics directly to consumers (Obermeier et al. 2022; Trujillo Falcón et al. 2022). Stumpf and Gerard (2021) concluded the following:

> Perhaps the most critical short-term need to move TIM [threat-in-motion, a type of PHI forecast graphic] forward is to establish optimal data formats as well as dissemination and notification modalities. Particular focus should be made on systems such as the Integrated Public Alert and Warning System, the Emergency Alert System, the Wireless Emergency Alert system, and NOAA Weather Radio, for television, radio, Internet, and mobile technology, in order to meet the needs of those end users and assure that public receipt of warnings remains whole (Stumpf and Gerard 2021, p. 642).

Recent studies have likewise discussed the possible use of mobile devices for presenting PHI forecast graphics to consumers (Ding and Millet 2020; Miran et al. 2018, 2020; Wehde et al. 2021). Obermeier et al. (2022) cautioned, however, "Any eventual operational warning product must work in alignment with PHI, lest serious communication issues emerge for end users and potentially the public" (Obermeier et al. 2022, p. 961). Indeed, inclusion of PHI forecast graphics within WEA messages appears to contradict aspects of the warning response model (Mileti and Sorenson 1990), which maintains that effective warning messages must be clear and consistent. Nevertheless, Miran et al. (2020), argued that "probabilistic weather information should be provided to the public in a way that they can make decisions on taking protective action based upon their own thresholds rather than defining thresholds by the officials and recommending people to take protective action at certain thresholds" (Miran et al. 2020, p. 8). Thus, debate exists concerning whether and how PHI forecast graphics should be shared via mobile devices with consumers, if at all.

Drawing together WEA system and FACETs developments, the goal of this study was to understand how PHI-enhanced WEA360 messages might influence people's risk perception and protective action decision-making for a tornado. To gain this understanding, this study investigated what groups of consumers think and say when shown mock PHI-enhanced WEA360 messages for tornado. Specifically, this study collected and analyzed qualitative data gathered from a combination of questionnaire (open-ended and scaling questions) and focus group interview methods conducted in collaboration with 31 community members in Denver, Colorado. The mock PHI-enhanced WEA360 messages used for the study combined two elements: (i) a slightly modified version of NWS's existing WEA360 message template for tornado and (ii) PHI forecast graphics that resembled those used in prior experimental research (Shivers-Williams et al. 2021). While such messages are not currently permitted via the WEA system, the FCC's 30 March 2023, Further Notice of Proposed Rulemaking indicated that in the future it may be possible for "alerting authorities to send links to location-aware maps in WEAs, allowing consumers see where they are relevant to the emergency situation" (FCC 2023, p. 1). While currently disallowed by the FCC, the issuance of PHI-enhanced WEA360 messages in the future may be technically possible. This study indicates why any such issuance through the WEA system, or other mobile device-based alerting systems, should be carefully considered, if not completely avoided.

The study begins with a narrative literature review in section 2, followed by a description of the qualitative data collection and analysis procedures in section 3. Section 4 presents the results of this analysis grouped by the major themes that help account for participants' risk perceptions and protective action decision-making. Section 5 discusses the limitations of this study, and section 6 provides a summary and conclusion.

2. Literature review

In this section, we review literature in two areas that informed this study. First, we highlight recent research that engages people's cognitive and behavioral responses to WEA messages. Second, we discuss studies of PHI that have focused on the interpretations of "downstream" consumers rather than "upstream" forecasters and meteorologists. By downstream, we mean research focused mostly on public interpretation and response to PHI-enhanced text and forecast graphics, rather than upstream research focused mostly on technical issues or forecaster interpretation and use of PHI products. The narrative literature review was assembled using databases including Academic Search Premier Plus and Google Scholar. In addition, the presentation database of the annual meeting of the American Meteorological Society (AMS) proved exceptionally useful. Keywords used in searches included "wireless emergency alerts" and "probabilistic hazard information," as well as combinations of these terms (both exact and nonexact matches) with words including "tornado," "forecast," and "response."

a. Public responses to WEA messages

Researchers have outlined a theoretical and applied communication research agenda for character-limited ("terse") public warning messages. This agenda involves studying (i) how hazard-related information can best be communicated in short messages, (ii) how a map or other location-related information might be included, (iii) how messages can be configured and disseminated to minimize delay time and maximize personalization, and (iv) how individual and contextual factors influence mobile public warning message reception, comprehension, and response (Bean et al. 2015; Kuligowski and Doermann 2018; National Academies of Sciences Engineering and Medicine 2018; Sutton and Kuligowski 2019).

While use of the WEA system has historically outpaced research concerning its benefits, limitations, and efficacy (Bean 2019), this situation is now changing (Cain et al. 2021; LaForce and Bright 2021a,b; Lee and You 2021; Ling and Oppegaard 2021). Researchers have found that subtle differences in the way that WEA messages are written can increase or decrease recipients' understanding, message believability, and personalization (the assessment that a disaster warning is aimed at oneself) of their content (Wood et al. 2018). For example, instead of including only county-level geography in WEA message text, including the words "in this area" can increase a message recipient's perception of personal risk (Bean et al. 2014). These subtle differences can also influence people's use of critical network resources, which can impact emergency service provision (Lambropoulos et al. 2021).

Research concerning the use of the WEA system to warn people facing an emergency has typically focused on correlations between WEA message attributes (i.e., source, hazard, guidance, time frame, location, style, and map and/or URL inclusion) and recipients' interpretations (comprehension and belief), personalization, and behavioral intentions and actions, that is, protective action decision-making and response (Bean et al. 2016; Doermann et al. 2020; Kim et al. 2022; Kuligowski and Doermann 2018; Liu et al. 2017; Sutton and Kuligowski 2019; Wood et al. 2018). Efforts to improve public response to COVID-19 health guidance and requirements has accelerated this "message optimization" focus within WEA system research (Bean et al. 2022; Lee and You 2021).

The message optimization paradigm used in the mobile public alert and warning research arena stems from the work of Dennis Mileti and colleagues (Mileti and Sorenson 1990; Mileti and Peek 2000). Mileti's warning response model (also known as the "hear-confirm-understand-decide-respond" model) has been foundational in the field of public warning research (Bean 2019; Sellnow and Seeger 2021). The model maintains that effective public warning messages include five content elements: source, hazard, guidance, location, and time (Mileti and Sorenson 1990). Effective messages are also accurate, clear, certain, specific, and internally consistent (Mileti and Peek 2000). A substantial portion of the mobile public warning research literature has been produced by Dr. Mileti's advisees and collaborators. These researchers have promoted the message optimization paradigm and the hear-confirm-understand-decide-respond model, often combining it with Lindell and Perry's (2012) protective action decision model (PADM).

Sutton et al. (2021) found that inclusion of enhanced protective action guidance in the message text, weather graphic, or both elicited increased understanding of the message. The researchers claimed, "The inclusion of protective action content in messages, whether via text, infographic, or a combination, resulted in increased ability to make decisions about the message, as well as increased self-efficacy and response-efficacy among participants" (Sutton et al. 2021, p. 24). Importantly, the researchers speculated that in the case of mobile public warning messages, which typically are delivered to large populations, "even small 'effects' can make a substantive difference for large numbers of individuals in terms of reduced numbers of deaths and injuries, reduced economic losses, and improved quality of life" (Sutton et al. 2021, p. 25). Sutton et al.'s (2021) study suggests that "enhanced protective action guidance," can be included in mobile public warning messages in ways that make a difference.

Wood et al. (2018) argued that when people are provided with warning information about an imminent threat, they generally seek additional and confirming information before they act—a process called "milling." Ling and Oppegaard (2021) identified this social behavior in their study of public responses to Hawaii's false missile alert in 2018, finding that 73% of WEA message recipients tried to confirm the message using another source. Often, these sources included people with whom the message recipient had strong interpersonal ties. In general, milling behavior delays the onset of protective action, which can be deadly in a fast-moving emergency. Information seeking and confirmation "erodes the advantage provided by warnings for imminent threats" (Wood et al. 2018, p. 536).

Wood et al. (2018) used emergent norm theory (ENT) to describe how new group norms can emerge in response to a hazard through interaction with others: "As new norms emerge, group members continue to act as individuals, but choose similar behaviors for shared reasons" (Wood et al. 2018, p. 538). Importantly, the researchers note that "[m]illing does not require a gathering of confused people . . . it is not the physical act of milling, but rather the psychological state of yearning for direction that yields the emergence of new norms" (Wood et al. 2018, p. 538). ENT predicts that the initial reaction to a PHI-enhanced WEA360 message will be to yearn for more information and instruction. Message recipients may interact with others to form shared understanding. Protective action may follow (or not) from "whatever socially constructed definitions might result from milling" (Wood et al. 2018, p. 538). Wood et al. concluded that even well-written warning messages might not eliminate milling because norms must still be socially sanctioned through interaction with others [for a discussion of ENT vis-à-vis earthquake early warning, see McBride et al. (2021)]. The PHI paradigm, however, has so far not considered the influence of milling behavior, nor emphasized protective action guidance, as we discuss next.

b. *Public responses to PHI forecast graphics*

PHI forecast graphics are a type of phenomenon-based warning; that is, PHI forecast graphics provide information about a weather phenomenon (e.g., tornado), but information about likely *impacts* (e.g., injury or death) and recommended *protective actions* (e.g., take shelter) are absent. Instead, PHI forecast graphics are based on the assumption that people perceiving higher probabilities of impact will understand their increased personal risk and therefore be more likely to take appropriate protective action (Miran et al. 2019). However, it is important to underscore upfront that "one cannot assume a priori, that additional hazard [e.g., PHI] or impact-based messaging will increase risk perceptions and motivate additional people at high risk to take protective actions" (Morss et al. 2018, p. 56). Numerous studies (several discussed below) indicate that improved risk perception and improved protective action decision-making do not necessarily proceed in lockstep.

For example, Shivers-Williams et al. (2021) examined the interplay between tornado warnings and PHI on the public's protective decision-making using a mixed-method approach involving 3003 online participants nationwide. The study used experimental PHI graphics to assess the interactive effects of warning type, hazard type, storm probability, and labeling scheme on protective-action decision-making. Participants viewed four graphics: a high probability tornado, a medium probability tornado, a high probability severe thunderstorm, and a medium probability severe thunderstorm. Participants were more likely to take protective action when responding to a tornado than when responding to a severe thunderstorm. They were less likely to take action when responding to PHI graphics with no warnings. They responded positively to the PHI graphics but noted that they would still want a textual warning message to accompany the graphics. Shivers-Williams et al. found a complex interplay among warning text, graphic displays, and protective action decisions. The bottom-line, however, was that PHI-enhanced warnings did not demonstrate a linear, causal effect on people's decision-making.

Childs et al. (2021) also explored public perception of different graphical PHI formats, as well as probabilities in general. The researchers conducted 36 semistructured interviews with members of the public in Tuscaloosa, Alabama to gauge their beliefs and responses to deterministic, red-color probabilistic, multicolor probabilistic, and textual probabilistic tornado warning graphics as part of the VORTEX-Southeast (VORTEX-SE) research program. A deterministic polygon has a single boundary and does not differentiate areas of varying risk within its boundaries (Jon et al. 2018). The three graphics depicted a geographical basemap with a blue dot indicating a participant's assumed location near the northeastern part of the polygon in a 30% probability zone. The textual information stated a 30% chance of a tornado at their specific location with different probabilities at other locations. Each interview began by successively showing participants the four graphics and asking the following questions: (i) What do you think this information means? (ii) What would you do if you received this information and why? The researchers found that the traditional, deterministic polygon prompted more people to take immediate protective action than red-color probabilistic or textual warnings. Importantly, inclusion of numerical probabilities appeared to allow participants to "game" their personal risk in ways that were not possible with the deterministic polygon. Childs et al. found that people's personal risk assessment processes could be independent of PHI format.

Much of the published work in the PHI area has been produced by Seyed Miran and colleagues, who have investigated (i) graphical designs for PHI that can increase users' perception, improve their interpretation, and elicit an optimal response to tornado threats (Miran et al. 2017a); (ii) PHI images in comparison with traditional polygon images for tornado warning (Miran et al. 2017b); (iii) how PHI encourages protective action intentions for tornado threats (Miran et al. 2018); (iv) people's thresholds for decision-making using PHI (Miran et al. 2020); (v) how PHI-related behavioral intentions compare with intentions influenced by traditional polygon-based warnings (Miran et al. 2019); and (vi) PHI's influence on people's levels of fear and protective actions (Miran et al. 2021). Briefly, Miran et al. (2017a) concluded that "four-color" PHI without an underlying radar image produced the best outcomes, as did a "five-color" PHI visual when compared with a deterministic polygon. Miran et al. (2018, 2019) found that in certain tornado proximity and likelihood scenarios, a PHI visual compelled more timely and appropriate protective action than a traditional polygon warning. Yet, Miran et al. (2020) found that accurately predicting when someone would act based on a PHI visual remained elusive. Nevertheless, using an experimental design, Miran et al. (2021) concluded that "providing the likelihood of occurrence of a particular threat through the color-coded swaths, as compared with the deterministic polygons, increased the level of fear and protective action in general" (Miran et al. 2021, 13–14). This finding contrasts with Shivers-Williams and Klockow-McClain's (2020) study, discussed below, which found that the likelihood of taking preparatory action increased with categorical forecasts relative to probabilistic forecast information.

Miran and colleagues' work raises two key issues that the present study addresses (i) the need to elicit the *diversity* and *complexity* of people's actual understandings and responses to PHI and (ii) the need to understand the connection between PHI's emphasis on shifting probabilities on one hand, and the WEA system's requirements for high levels of certainty before message issuance on the other. Miran and colleagues' PHI research overwhelming focuses on *individual* sensemaking and response, but we know from research on public alert and warning that protective action decision-making and response is a highly *social* behavior fraught with complexity, ambiguity, and contradiction (Mileti and Sorensen 1990; Wood et al. 2018). Miran and colleagues' PHI research also highlights the desirability of probability matching and urgency matching, which should, in theory, lead to more appropriate public decision-making and response. However, PHI potentially increases the ambiguity of warning messages as content moves from binary instruction (e.g., take shelter or evacuate) to more subjective and nuanced assessments of personal risk.

Shivers-Williams and Klockow-McClain (2020) explored the potential trade-offs between showing people a probability value that is relatively large, or a value that reflects an event that is likely to occur nearby. In other words, the researchers investigated which factor had more influence on protective action decision-making: higher probability of occurrence or nearness. Participants in the study responded to a series of decision scenarios consisting of severe weather probability forecasts and fictitious maps, with "likelihood of taking preparatory action" serving as the dependent variable (Shivers-Williams and Klockow-McClain 2020, p. 8). Importantly, the researchers found that "categorical forecasts encouraged taking preparatory action, on average, while probabilistic forecast information, even when coupled with categorical forecast information, seemed to lessen the propensity to take preparatory action by comparison" (Shivers-Williams and Klockow-McClain 2020, 15–16). This finding has implications for PHI-enhanced weather warning because it suggests, contrary to other studies, that categorical warning (e.g., high, moderate, enhanced, slight, and marginal) may produce better protective action outcomes than probabilistic warning. In short, the finding seemingly challenges a key premise of the PHI paradigm. However, the scale of the warning area (city, county, region, etc.), and the message's textual and visual content, influenced outcomes in complex and significant ways.

Similarly, Jon et al. (2018) investigated whether a gradient polygon (similar to PHI swaths) would be useful in spurring appropriate protective action in response to a tornado hazard. Jon et al. noted that prior research has established a "centroid effect" in response to visual polygons; that is, people generally perceive a tornado's strike probability to be higher at the center of the polygon, even though NWS guidance "implies that all locations within the polygon are equally likely to be struck" (Jon et al. 2018, p. 4). Likewise, prior research has established an "edge effect," whereby people who believe themselves to be located just outside the edge of the polygon make similar risk assessments to those that people just inside the edge make—even though there is negligible risk of tornado strike outside the polygon. In reviewing much of the same literature included in this study, Jon et al. concluded that "existing research has shown that a probabilistic polygon-only display is superior to conventional deterministic polygon-only display in producing increases in [strike] judgments at the near edge of the polygon and, thus, producing expected protective actions that are more consistent with NWS guidance" (Jon et al. 2018, p. 4). The question remained, however, whether NWS should superimpose a probabilistic swath, rather than a deterministic polygon, onto its radar display forecasts. Through experiments (mostly with college students), Jon et al. (2018) found that decision-making outcomes were similar for displays of conventional radar images and displays of gradient polygons (PHI-type images). In short, they concluded that "it makes little difference what type of polygon the NWS superimposes onto radar displays" (Jon et al. 2018, p. 13). Additionally, the study confirmed people's tendency to perceive (i) higher risk at the polygon's centroid than at its near edge, (ii) less risk moving out from the centroid, and (iii) much less risk just outside the edge of the polygon (despite risk being negligible).

Importantly for the present study, Jon et al. (2018) found that participants had no general tendency to prefer any particular protective action anywhere inside the polygon, regardless of the type of display viewed. PADM (Lindell and Perry 2012) predicts that participants should have indicated a greater need to shelter promptly as strike probability increased. Jon et al. (2018) stated that a possible explanation for their counterintuitive result lay in participants' generally low levels of experience with tornadoes.

Joslyn et al. (2020) investigated the impact of color-coded probabilistic tornado warnings on risk perceptions and responses via experiments and interviews. In an experimental study, participants assessed four different visuals for a tornado, two of which were PHI-based. Importantly, participants were told to imagine that the images would appear on their mobile devices, but the images were not formatted with the WEA system in mind. The accuracy of participants' perceptions of tornado probability improved with the PHI-based visual, relative to a deterministic polygon. Moreover, participants who were given the deterministic polygon visual chose to shelter less frequently at higher probabilities and chose to shelter more frequently at lower probabilities in comparison with participants who were provided the PHI-based visual. The PHI-based visual also resulted in participants trusting the forecast more, but there was some reported confusion regarding the differences between tornado likelihood and severity. The question of whether PHI-enhanced messages improve protective action decision-making has thus produced divergent findings, with Jon et al.'s (2018) study suggesting minimal improvement and Joslyn et al.'s (2020) study suggesting major improvement.

Demuth et al. (2020) used interviews to delve into the findings of Joslyn et al.'s (2020) experimental study of color-coded probabilistic tornado warnings. The interview protocol mimicked, in many ways, Joslyn et al.'s experimental design, with images of deterministic polygons, gradient polygons, and numerical probabilities of occurrence. Participants were shown their location at different points inside and outside the polygon and asked questions related to three dependent variables: likelihood of tornado strike at their location, what they would do if they were in their home, and how much damage they think the tornado would cause. Using interviews, the researchers were able to probe for why participants made the interpretations and decisions that they did. Presented with textual information only, participants assessed the likelihood of tornado strike at their location in ways that paralleled the (presumed) actual probability. Participants' estimates of likelihood surged well beyond actual probability when shown deterministic polygons and gradient polygons. Similar to Jon et al.'s (2018) findings, Demuth et al.'s (2020) found that high means for intended action reported in experiments occasionally mask variability in people's sensemaking processes. Consistent with prior research, some people nearest the tornado hazard will still not take recommended protective actions despite clear indications of personal risk (Jauernic and Van Den Broeke 2016; Walters et al. 2019, 2020).

In sum, the research record suggests that while stakeholders may anticipate PHI forecast graphics will improve people's accuracy of risk perception, those improvements might not necessarily lead to substantial increases in the number of people taking appropriate protective action. The present study sought to investigate the issue further by combining WEA360 messages with PHI forecast graphics and then qualitatively assessing people's responses to them. The research question guiding this study was: How might PHI-enhanced WEA360 messages influence people's risk perception and protective action decision-making?

3. Method

This study unfolded in four phases: stimuli development, recruitment, community focus groups, and a meeting with officials.

a. Stimuli development

To create the mock PHI-enhanced WEA360 messages, the authors copied the existing WEA360 template for tornado (see https://www.weather.gov/wrn/wea360) and modified the end of the message to better account for the inclusion of the PHI forecast graphic. WEA360 message templates include information about the source, hazard, location, time, and recommended protective action. PHI forecast graphics can be presented in different ways, but they generally include an underlying map, a PHI plume that represents the entire area at risk (the outer edge of the largest PHI swath), different color-coded PHI swaths within the plume that represent different probabilities of tornado strike, and a legend that indicates the numerical probabilities of the different PHI swaths. The PHI forecast graphic used in the present study was derived from the Shivers-Williams et al. (2021) study (see Fig. A5 in the appendix). Following recent experimental studies of PHI, a "blue dot" was included in the mock messages to indicate the recipient's location (Demuth et al. 2020). This blue dot was meant to mimic the marker used in some mapping applications (e.g., Google Maps). Although current WEA system technology does not indicate one's location on a map in this way, the FCC has recently issued a Further Notice of Proposed Rulemaking that would allow it in the future (FCC 2023).

To mimic a dynamic PHI forecast graphic received on a mobile device, the authors developed a set of static images that could, in theory, be issued via the WEA system roughly 10 min apart. The entire PHI plume in the mock messages represented approximately 1 h of warning time. Images received in this way mimic the type of weather warning messages available via some types of weather apps (e.g., RadarScope). Following an NWS official's recommendation, the background map used for this study was derived from the RadarScope app developed by Base Velocity, LLC. In theory, such images could be added to a website accessible via an embedded reference within a WEA message (still shots of PHI via embedded multimedia may be on the horizon, according to the FCC). It would have been desirable to include a dynamic gif of PHI, but this technology is not easily embedded in a WEA message, and what elements of a PHI gif would maximize outcomes is uncertain.

The authors used Adobe Photoshop to create the mock PHI forecast graphics. First, the RadarScope app was used to generate a map of the downtown Denver area, and the app's zoom feature was used to obtain the most detailed map possible for the ~50% and ~90% probability of impact maps, that is, no further zoom was possible, and the map included the names of cities and towns in the Denver metropolitan area. Because the scale of the ~10% probability of impact image and map necessarily differed from the other two images, the authors decided that the ~10% message/visual would be presented separately to participants to avoid having them perceive continuity among all three messages. To repeat, the entire PHI plume in each mock message (all of the PHI swaths) represented approximately 1 h of warning time end to end. Next, the authors copied an NWS radar image of the 10 August 2020, derecho, a type of thunderstorm (https://www.weather.gov/dmx/2020derecho), overlaying a portion of that image on top of the RadarScope map to create a sense of extreme weather. Third, the authors recreated the PHI plume used in the Shivers-Williams et al. (2021) study, as well as copied the probability legend from that study [participants slightly preferred it in comparison with the other legend types that Shivers-Williams et al. (2021) tested]. Also, blue dot resembling a Google Maps location marker was added to indicate the recipient's location at the Auraria Higher Education Center (where the sessions occurred). The authors generated the messages using the WEA360 tornado template (provided in English and Spanish) but replaced the "check media" guidance with impact-based historical context and probability of tornado strike information. This explanatory text reflected Ripberger et al. (2020) and Ding and Millet (2020) PHI-related message writing guidance. As noted, the Spanish versions of the mock messages omitted accent marks to match current WEA360 templates [for a discussion of potential problems with this omission, see Trujillo-Falcón et al. (2021)]. The study's mock messages were presented to members of several NOAA/NWS units for feedback. Following discussion and refinement, the final mock messages used in the study in both English and Spanish were produced (see the examples in appendix Figs. A1–A4).

b. Recruitment

In the next phase, the authors recruited 31 community focus group participants via the Denver Craigslist website. An incentive ($150 supermarket gift card) was provided to each participant. The authors sought to recruit community members who represented the demographic characteristics of Denver County. According to the U.S. Census (https://www.census.gov/quickfacts/fact/table/denvercountycolorado/PST045222, accessed 10 December 2020), Denver County has a diverse population that is reported as 29.3% Hispanic, 9.8% African American, 4.1% Asian, and 1.7% Native American. The 31 participants' reported racial and ethnic identities were 23% Hispanic, 6% African American, 6% multiethnic, and 6% preferred not to disclose. The remainder, 61%, reported as White. Regarding educational attainment, 88% of the population in the metropolitan area have a high school diploma, and 49.4% have at least a bachelor's degree. Of the 31 participants, 52% reported holding a bachelor's degree or higher. 11.9% of Denver County residents are over 65, while only 6% of participants were 61 or older: 13% were 18–25, 55% were 26–35, 10% were 36–45, and 16% were 46–60. Males total 50.1%, and females total 49.9% of the population in Denver, while the participants totaled 58% and 42%, respectively. The median household income in Denver is $68,592, and 90% of the participants reported annual income at or below $59,000. The authors sought to include residents with Spanish-language proficiency in the groups because WEA messages can be issued in both English and Spanish. Participants were asked to self-assess their Spanish reading proficiency, and 29% of participants reported "medium" or "high" levels. While not an exact mirror, the participants' demographic characteristics reasonably resembled those of Denver County for the purposes of the study. Destructive tornadoes have struck Denver before in 1981, 1988, 2013, and 2019 (Smith and Horner 2021). Each group of participants contained a mix of people with differing levels of tornado knowledge and experience. (The demographic characteristics of the participants are presented in Table A1 in the appendix.)

c. Community focus groups

The next phase involved six community focus groups that combined questionnaires and interviews based on a tornado warning scenario. In testing the public's response to receiving severe flood warnings using simulated cell broadcast, Smith et al. (2022) combined similar methods. In the present study, each focus group saw four mock messages (described below). Each participant was presented with a mock WEA360 message and embedded PHI forecast graphic on a single piece of laminated paper with the image roughly the same size and shape of a typical smartphone. To increase the realism of the scenario, participants were asked to imagine that the conference room in which they were meeting was possibly in the path of the tornado under discussion. The conference room was located on the fourth floor of a five-story building, with a floor-to-ceiling glass wall on one side of the room. Once presented with the scenario and the first mock message, the participants were asked to silently complete a questionnaire to capture their initial, individual interpretations and responses to the first message: ~50% probability of impact. The questions were drawn from Bean et al.'s (2014) study of WEA messages and embedded maps and included the following:

1) On a scale of 1 to 7 (with 1 being very low and 7 being very high), please rate your understanding of the message. Is there anything about this message that you are unsure about? If so, please explain.
2) On a scale of 1 to 7 (with 1 being very low and 7 being very high), please rate your belief in the message. Is there anything about this message that would question your belief in it? If so, please explain.
3) On a scale of 1 to 7 (with 1 being very low and 7 being very high), please rate your personal risk of tornado impact based on the message.
4) What action would you take in response to this message?
5) On a scale of 1 to 7 (with 1 being very low and 7 being very high), please rate how *confident* you are that your decision/action in question 4 is appropriate for the risk that you face.
6) Please explain how valuable this message would be to you in an actual tornado emergency.

Following participants' completion of the questionnaire, they were asked to account for their responses to the mock message. A similar "think-out-loud" approach (Ericsson and Simon 1993) has been used in the study of mock WEA and Twitter messages (Bean et al. 2014), cell broadcast messages in general (Smith et al. 2022), tsunami warning messages (Sutton and Woods 2016), and tornado warnings (Sutton and Fischer 2020). Specifically, the participants were asked what action they had decided to take based on their written response to question 4 in the questionnaire. Participants were subsequently asked to explain *why* they had decided to take that action, and after that were asked, "Based on this message, what do you think we should do?" This question was designed to spark social interaction within the group in response to the message. The participants described their thoughts and feelings as they interpreted, reread, questioned, or puzzled over the mock WEA360 message content and PHI forecast graphic. After interacting, a second questionnaire that included the same questions as the first questionnaire (omitting question 6 to save time and avoid redundancy) was given to the participants to capture whether any changes in responses had occurred.

For all six community focus group sessions, each group saw four mock messages. The ~50% probability of impact message was presented first, followed by the ~90% message. We wanted to learn what interpretations and decisions participants would first make in a situation of ambiguity, hence the initial presentation of the ~50% probability of impact message. Then, after explaining the third message was not connected to the prior two, the ~10% message was distributed. The ~90% probability of impact Spanish-language message was then shared with the groups. In other words, each participant viewed four mock messages. Groups 1 (four participants), 3 (three participants), and 5 (six participants) had impact-based historical context included in all four of their messages, while groups 2 (five participants), 4 (seven participants), and 6 (six participants) had no impact-based historical context included in their messages.

Community session transcripts were generated by Landmark, a professional transcription company. The first author reviewed/corrected transcripts while watching/listening to the video recordings. The transcripts were analyzed using an iterative approach (Tracy 2019). An iterative approach differs from a grounded theory approach (Strauss and Corbin 1997) in that grounded studies often delay the literature review until after data are collected to help ensure an inductive examination of the data. An iterative approach, by contrast, alternates between emergent readings of the data and use of existing models and theories; it reflects the interests, priorities, and theories the researcher brings to the data (Tracy 2019). An iterative approach was used to (i) generate first-level codes to characterize participants' utterances, (ii) group utterances together using the study's research questions as secondary codes, and (iii) identify concepts from the research literature (reviewed above) that could account for the patterned utterances.

For this study, the goal was to describe and explain the *multiple perspectives* that arose among participants—perspectives constituted by unique personal histories and diverse social contexts. To ensure quality, the authors accounted for the range of perspectives offered (noting "deviant cases"), used plentiful examples of transcript segments for illustration, triangulated between different data types (utterances and written responses to two questionnaires), and asked participants to validate the legitimacy of analytic interpretations (O'Connor and Joffe 2020; Phillips and Hardy 2002). This last step was done in two ways: during the sessions themselves (with the researcher asking the group if an overarching interpretation of their responses seemed reasonable) and prior to the meeting with officials (see the next section) where themes from the focus group sessions were provided to two community member participants and discussed with them.

d. *Meeting with officials*

In the last phase, the authors confirmed two of the community members (via e-mail invitation to participants) to join a final meeting with two NWS officials and a Denver emergency management official. These five stakeholders were brought together in collaborative partnership to discuss weather warning (Takenouchi et al. 2017). This final meeting involved discussion and consideration of themes emerging from the community focus group sessions: probability thresholds, urgency matching, time of arrival, impact-based historical context, and Spanish-language messages. The purpose of this meeting was to help pinpoint opportunities and obstacles involved in combining WEA360 messages and PHI forecast graphics in ways that respond to local concerns. The participants in the final meeting were asked to both collaborate and reflect the interests of their respective groups as they discussed the usefulness and feasibility of various configurations of PHI and warning approaches. The coded transcript from this final meeting involved the same analytical approach described in phase 3 and incorporated into the thematic analysis.

4. Thematic analysis and discussion

Community focus group participants' comments revealed challenges of combining WEA360 messages with PHI forecast graphics. In a nutshell, the overarching theme uncovered was that WEA360 messages that include PHI forecast graphics risk violating a key principle of effective public alert and warning: internal consistency (Mileti and Sorenson 1990; Mileti and Peek 2000). Critically, perceived mismatches between the WEA360 message's instructional guidance and the PHI forecast graphic raised the question of what probability threshold *should* compel immediate protective action. At ~90% probability of impact, all 31 community focus group participants stated their intention to immediately take shelter. At ~50% probability of impact, 19 participants (61%) initially indicated their intention to immediately take shelter, falling to 16 (51%) after group discussion. At ~10% probability of impact, almost no participants intended to immediately take shelter. These results suggest a correlation between probability of tornado impact and message recipients' intention to take shelter, resembling the well-established "proximity heuristic" (Jon et al. 2018; Lindell and Earle 1983). This heuristic maintains that "perceived risk decreases with distance from the expected impact location (Jon et al. 2018, p. 5). Yet, at what probability level a PHI-enhanced WEA360 message recipient *should* take shelter remains a question whose answer will necessarily influence one's interpretation of this study's findings.

The underlying assumption of the PHI paradigm is that people who are presented with forecast graphics depicting higher probabilities of impact will understand their increased personal risk and therefore be more likely to take appropriate protective action. Likewise, people who are presented with lower probabilities of impact will presumably understand their decreased personal risk and therefore avoid unnecessary protective action. This study's results support the reasonableness of that underlying assumption. Prior experimental studies have demonstrated that PHI can improve risk perception (Joslyn et al. 2020; Miran et al. 2018, 2019, 2020, 2021; Shivers-Williams et al. 2021); yet these studies have not established consistent improvements in protective action intention. This study's results likewise align with the outcomes of these earlier experimental studies. Patterns in message-related and nonmessage related factors that emerged from analysis of participants' comments help to account for these results. Three major themes arising from this analysis are presented next: understanding and belief; language expectancy and urgency matching; and milling.

a. *Understanding and belief*

The community focus groups revealed differences in people's understanding of PHI forecast graphics' indication of the time of arrival of the tornado. Consistent with previous studies (Gillespie et al. 2022), the authors constructed the mock PHI forecast graphics used in the project (in collaboration with NOAA/NWS officials) to represent approximately 1 h of time, end-to-end, for the entire PHI plume, that is, the outer edge of the entire group of PHI swaths. Asking one set of seven community focus group participants how much time they thought the full plume of PHI swaths represented, answers ranged from "8 to 10 minutes" to "one hour." Three participants believed the plume to represent 20 to 25 min. Uncertainty about time of arrival confused and dismayed a few community focus group participants. One remarked, "[The message is] not specific about the length of time it's going to take to hit where I'm at. I have absolutely no idea, and especially if I'm new to Denver, and I just relocated." Another mentioned, "if there was somewhere up there [in the message] where you could get a link that would show you maybe time zones. Is this 10, 20, 30, et cetera [minutes to time of arrival]. Then you can guess what your probability is and how much time you have." Another participant noted:

> The 4:30 makes me wonder, because it's basically almost 4:30 right now. We're like, 'Oh, maybe we just got it late, so we're good.' That would be my question. Are we safe for 4:30, then? Just seek shelter for five minutes?

While some community focus group participants sought personalized time-of-arrival information, such information is currently beyond the WEA system's capability but may be possible with other types of multihazard apps (Dallo et al. 2022). The Common Alerting Protocol (CAP) that undergirds the WEA system includes an optional "onset" code that specifies the "expected time of the beginning of the subject event of the alert message." However, CAP has no capability to populate an alert message with information about an individual user's specific time of arrival for a given hazard. It may be possible, however, for officials to manually include general time frames in a message for recipients located inside a band of PHI swaths, such as, "A tornado may impact your location within the next 20 minutes." One NWS official expressed reluctance to use such language, however, and asked during the final meeting, "I guess my question too, then, is when you're talking about time progression, if your blue dot is out in the 20% [swath], are people anticipating that over time, their risk level will increase, or are you going to stay 20% for the next half hour or however long that warning is in effect?" Another NWS official asked during the final meeting:

> If your blue dot is out here at the edge, you're in just the edge of the 20% line, say 20 min from now, do you think your risk level, the probability, will that change for you? I guess what I'm asking, are you envisioning this whole thing shifting east, and 20 min from now, you might be in the orange level? Do you think you'll stay 20%?

The two community members participating in this session differed in their responses to the question, with one stating, "I think at 20%, lots of things could happen. I wouldn't guess. I would be paying attention." The other community member responded, "I think, for me, I would see that it's kind of dissipating. There's more area, so I'm kind of thinking to myself, that's not going to hit me." Thus, inclusion of PHI forecast graphics in tornado warning messages in the way done for this study raises the question of whether recipients will assume the risk of impact at their location is likely to intensify or dissipate over time. Time of arrival is a complicated issue that affects both PHI and dichotomous warning, but it appeared to be a critical factor in spurring protective action in the present study (see also Dallo et al. 2022).

Consider that within the "hear-confirm-understand-decide-respond" model (Mileti and Sorensen 1990), protective action decision-making is preceded by understanding the warning message and believing that the message is true, accurate, and meant for the recipient. The mock PHI-enhanced WEA360 message depicting a ~50% probability of impact used in this study elicited high levels of recipient understanding and belief (measured on a 1–7 scale), both of which are generally necessary for spurring appropriate protective action. The questionnaire used at the outset of the community focus groups asked participants, "Is there anything about this message that you are unsure about? If so, please explain" and "Is there anything about this message that would question your belief in it? If so, please explain." These questions were included to elicit (and pinpoint) uncertainties regarding the content or format of the messages. This initial questionnaire was distributed at the outset of the community focus groups before any discussion of the ~50% probability of impact message occurred.

Community focus group participants reported high levels of understanding (6.0 average) and belief (5.9 average) on a 7-point scale in response to the ~50% probability of impact message. For the six participants who rated their level of understanding at "5" or below, the meaning of the PHI swaths, location of the tornado, blue dot, and legend created confusion. For the nine participants who rated their level of belief at "5" or below, the wording of the message, its source, uncertainty about various message elements, prior tornado hazard experience and knowledge, and overalerting were cited as factors. Despite high levels of understanding and belief, however, only 19 of 31 community focus group participants indicated that they would heed the instructional guidance in the message and immediately take shelter. Nevertheless, participating NWS officials stated that a ~60% rate of intention to comply was encouraging. During the final meeting, one NWS official remarked, "60% is good, even with a yes/no warning. We'll take that." At ~90% probability of impact, every community focus group participant verbally indicated that they would take shelter. However, the ~90% probability of impact message was received after the ~50% probability of impact message had already been received and discussed. Message repetition (i.e., order effects) likely influenced participants' responses to the second message. It is unclear what the participants' responses would have been had they received the ~90% probability of impact at the outset of the community focus group sessions.

Mileti (2018) observed that people usually do not take immediate protective action in response to warning messages, even when levels of understanding and belief are high. This observation held true during the community focus groups. As predicted, nearly half of the participants claimed that they would first attempt to confirm the message via indicators in the physical, information, or social environment (e.g., looking outside at the sky, looking at one's mobile device for news, or observing what other people nearby are doing). Several participants indicated that they would seek to observe the weather around them, if possible, hoping to get a better sense of the urgency of the situation. As one participant remarked, "My thing is I'm a see-er, not a believer. If I see it, then I believe it. You've got to show me." Both understanding and belief were influenced by mismatches between the WEA360 message and the PHI forecast graphic, as discussed next.

b. *Language expectancy and urgency matching*

PHI's probabilistic foundations potentially increase the ambiguity of warning message content as recipients move from binary responses (take shelter or not) to more nuanced assessments of personal risk. CAP requires that WEA message writers specify an event's urgency (the time available to prepare), severity (the intensity of impact), and certainty (confidence in the observation or prediction of the event). Only the highest levels of these categories permit WEA message issuance. WEA message issuance requires a CAP "severity" code to be either "Extreme" ("extraordinary threat to life or property") or "Severe" ("significant threat to life or property"). CAP also requires officials to specify the "urgency" of the situation, but only the "Immediate" code seems appropriate for tornado threats, as the next available option, "Expected," indicates that "Responsive action SHOULD be taken soon (within next hour)." One hour is too far in the future for people located in probability swaths nearest a tornado. Therefore, to align PHI-enhanced WEA360 messages with CAP requirements, NOAA/NWS would need to identify how different probability thresholds could be mapped to "Severe" and "Extreme" CAP severity codes.

One approach might be to identify as "Severe" or "Extreme" *any* geographic area falling within the entirety of the PHI plume. Doing so would ensure that people located within all probability of impact swaths (from 1% to 100%) would receive the warning message. This approach risks triggering complaints of overalerting from some message recipients located in low probability swaths. Yet, most participants agreed that they would still want to receive warning messages despite a low probability of impact. One participant explained, "The ranges that aren't close to where I'm at, but I might know somebody in that area or something. That's why I would like the alert—to let them know if the percentage is crazy high." Another participant, however, expressed irritation at receiving the ~10% probability of impact message:

> Like you [another participant] were saying, it [the text] just makes it sound so extreme, like it's a live or die type situation. I feel like we have so much stuff on our phones these days too, whether it's something like this, or any kind of other warning or AMBER Alerts or just regular text messages, or social media notifications. At some point you just—they can all add up and you just kind of swipe it all away.

Thus, supporting Miran et al.'s (2019) findings, adequately matching the urgency of the textual elements of the message to the PHI forecast graphic's indication of probability of impact emerged as a key concern. Community focus group participants noted that the sense of urgency in the language used in the textual portion of the message needed to better match the visual probability of impact. In other words, participants expect the language in the text to better match the PHI forecast graphic. One participant stated, "[I]f they are using a [software] program with the legend, you should be able to have different messages that correspond to the probability on the legend." Some message recipients in low probability swaths who receive instructions to take shelter "now" may question the accuracy or trustworthiness of the message. One participant remarked:

> From the language, from the text, it makes it sound like there is imminent danger if it is to be trusted. I'm not saying it's not trustworthy, but, if it's to be trusted, then—yeah—there's imminent danger headed your way, but then you look at the map [~10% probability of impact], and it's like, "Well?"

Effectively including PHI forecast graphics in WEA360 messages would require rewriting WEA360 text to better match each probability of impact indicated by each swath in the PHI forecast graphic. One participant noted:

> If it's just the same messaging every time, with all these different risks, different areas and all of that, then maybe it's easier to dismiss that [the warning]. I do not know how difficult it is to get one of these notifications based on an even more accurate measurement of your location, to be like, "Your risk is 0 to 20 percent."

The complexities of tailoring WEA360 language to each probability swath in a fast-moving, uncertain situation appear insurmountable with current technology and warning practices. Such precise textual and visual alignment does not appear viable within CAP, as finer distinctions would need to be made between "urgency" values of "Immediate" and "Expected." Moreover, the meeting with NWS and Denver OEM officials revealed that there is no risk threshold at which NWS forecasters *must* issue warnings. As one official explained, "Every forecast is going to be a little bit different with their [a forecaster's] personal threshold to be issuing a warning, but I think, on average, it's 50/50. If you're reasonably confident that it's just barely more likely there's going to be a tornado there versus not, we would issue it" (for a discussion of forecaster decision-making using PHI tools, see Trujillo-Falcón et al. 2022). For forecasters, there is significant benefit in moving from a dichotomous paradigm (warn or do not warn) to a PHI paradigm. This official continued, "Our stress levels would go down a little bit because then we would be providing that constant flow of information, not worrying so much about, okay, binary, yes, I'm going to send it now, or no, I'm not going to send it now because I'm just not confident." However, the theme of language expectancy and urgency matching suggests that reduced stress for message issuers may come at the expense of increased stress for message recipients. Those recipients are likely to turn to others to help them decide how to interpret and respond to the message, as discussed in the next section.

c. *Milling*

Reception of a warning message on one's mobile device within a group of people creates a situation in which social norms and expectations can influence interpretation and response. One's personal interpretation and response will be conditioned by the emergent norm that forms within the group (Wood et al. 2018). One participant characterized this dilemma:

> I think myself and anybody who's a transplant, we would probably feel out the room and look for anybody who even could remotely call themselves a native, because I wouldn't want to just be freaking out about every tornado, just as much as I do not want to be freaking out about every blizzard that comes my way this winter.

This participant did not want to appear either too hasty or unconcerned and would rely upon the group for cues. Whether a warning message is dichotomous or PHI-enhanced, within a group of people, warning response becomes a social phenomenon (Mileti 2018). Another participant explained:

> I'm more herd mentality. If everybody would stay here, I'd feel confident. If we were evacuating, I'd want to follow. If everyone decided to stay for it, do some research from there to make a decision . . . It just doesn't happen in Denver very often. I think I'd rather follow people who know better than me before making the call because I do not know.

The question for PHI stakeholders is whether and how PHI forecast graphics might become resources for sensemaking, interaction, and persuasion within the milling process. The present study examined participants' survey responses and utterances for clues about what PHI stakeholders can expect. When presented the ~50% probability of impact message at the outset of the community focus groups, participants silently completed a questionnaire that asked them what action they intended to take in response to the message. From the questionnaire, 19 of 31 participants indicated their intent to take shelter, usually in an interior room on the lowest level of the building. After participants had spent time sharing and discussing their reasons for their intended action, they again completed a questionnaire that asked them what action they intended to take in response to the same message. Analysis of the second questionnaire revealed that 16 of 31 participants intended to take shelter. The slight reduction in the number of people intending to take protective action can be attributed to the social meaning-making process ("milling") and emergent norms that arose within the different groups. Milling behavior and emergent norms offer both "good news" and "bad news" for PHI stakeholders. The good news is that group discussion produced improvements in understanding, belief, or intended protective action compliance for six participants (although some results were mixed). The bad news is that intended protective action also worsened for five participants overall.

Consistent with prior research, tornado experience inevitably shaped people's responses to PHI-enhanced WEA360 messages. Within a group, prior experience becomes a resource for sensemaking: one bound up with milling and message believability. Prior research suggests that verbally shared experiences can influence others' interpretations and responses (Bean et al. 2016). In this study, when looking for utterances that may have either reinforced or revised participants' intentions, both non-

PHI and PHI variables appeared influential. For some participants, previous false alarms or overalerting experiences generated skepticism of the message content and contributed to their reluctance to take immediate protective action. For others, however, certain message elements evoked memories of previous tornado emergencies, leading to their *increased* caution and willingness to act. For example, one participant stated:

> I totally hear where you're coming from [in not taking action], and I get it. I've ignored some of these things too in the past, but I do know the difference between a "warning" and a "watch," so I'd try and be like, "Okay. If they say it's a warning, and I've got an 80 to 100 percent chance of ending up in Kansas, then I'll probably go and seek shelter."

Determining exactly which utterances, like this one, produced changes in participants' perceptions is challenging but offers an enticing opportunity for future research.

In sum, we conclude that milling behavior and emergent norms contributed to both subtle improvements and degradations in some people's intentions to follow the protective action guidance included in the ~50% probability of impact message. For some participants, changes in intended protective action were mixed but usually slight. For others, no changes were apparent. The extent to which the PHI-enhanced content of the message *directly* shaped these outcomes could not be determined from the data collected. It is impossible to know with certainty exactly which utterances spurred which changes in people's perceptions. Nevertheless, as the example utterances above indicate, the influence of PHI-enhanced message content cannot be ruled out.

5. Limitations

This study's design necessarily created limitations. First, the "artificiality" of the community focus groups likely influenced the nature of the emergent sensemaking that occurred. In an actual emergency, people tend to seek warning confirmation with close ties (Ling and Oppegaard 2021), known acquaintances (Aguirre et al. 1998), and peers (Lindell et al. 2019). However, in absence of that possibility, people may interact with those not already known to them (Aguirre et al. 1998) and behave in prosocial ways (Bartolucci et al. 2021). A few participants noted that their responses would likely differ based upon who was with them when a warning message was received. Nevertheless, an "emergent group" (Drabek and McEntire 2003) may still form under the conditions used in the present study. Nevertheless, people often believe or say they will act in a certain way but act rather differently when faced with an actual emergency.

Second, the study relied on diverse participants but did not seek to correlate utterances with demographic factors in order to make claims about the influence or significance of those demographic factors. We know that demographic factors shape tornado warning response in general (Walters et al. 2019, 2020), but we do not know whether or how those factors *determine* social interactions in a milling situation. Investigating that relationship was outside the scope of this study. Likewise, there are multiple factors that influence tornado warning interpretation and response in general that were not systematically accounted for in this study that could have played an influential role in the communication exchanged during the community sessions: household composition, residence type, disaster myths, and fatalism, among others (Walters et al. 2019, 2020). Although the community focus groups' composition accounted for tornado knowledge and experience, again, the project did not seek to establish the *relative* influence of demographic, environmental, and experiential factors in producing an individual's level of knowledge or type of risk perception. This study instead focused on the language used within a diverse group of people and its influence on PHI interpretation and response, looking for patterns in participants' utterances and changes in their stated/recorded intentions.

Third, the study's setting in a sturdy, five-story building also likely shaped people's responses, as it was easy to imagine where/how to take shelter from an imminent tornado strike. The setting probably increased feelings of self-efficacy and response efficacy (Demuth et al. 2022). Studies of PHI accounting for a wider spectrum of tornado risk factors, perceptions, settings (including nocturnal scenarios), and actions should be conducted in the future.

Fourth, the mock PHI-enhanced WEA360 messages used in this study were presented on laminated paper. Although NWS officials confirmed that the mock messages approximated what actual PHI-enhanced messages might look like in the future (and their design was informed by prior studies), their content and format were speculative. It is possible that PHI-enhanced WEA360 messages could look different [see Lindell et al. (2022) for other types of graphics]. This study would also have benefited from evaluating each component of the mock messages (WEA360 text and PHI forecast graphics) separately with the same or different participants to assess whether the separate and combined components elicited different responses. Discussion of the Spanish-language messages suggested that while the absence of accent marks was noted, Spanish readers could still understand the content. We did not separate English and Spanish participants, however. Additionally, the mock messages used in this study were unable to reflect the requirements of access and functional needs populations. PHI-enhanced WEA360 messages could be especially problematic for people with visual challenges, a factor that potentially calls into question the PHI paradigm's reliance on color-coded swaths if related messages are shared with consumers (LaForce and Bright 2021a,b). Color distinctions, location markers, scales, legends, radar images, and other visual elements may present considerable challenges for access and functional needs populations. Careful attention to the needs of visually impaired populations must be considered in the future development and deployment of PHI.

6. Conclusions

This study engaged the question of how PHI forecast graphics included within WEA360 messages might influence recipients' understanding, belief, assessment of personal risk, and protective action decision-making. Supporting prior experimental research results (Joslyn et al. 2020; Miran et al. 2018, 2019, 2020, 2021; Shivers-Williams et al. 2021), this study found that inclusion of PHI forecast graphics within WEA360 messages generally enhanced participants' risk perception but did not spur protective action for nearly half of the message recipients when the probability of impact was ~50%. As the risk of impact increased, so did participants' intention to take protective action. As the risk of impact decreased, so did participants' intention to take protective action. However, changes in intention to take protective action were not linear or uniform among participants. Some participants used PHI forecast graphics to "game" their intended response in ways that could put themselves at greater risk.

PHI-enhanced weather warning via the WEA system could one day reduce overalerting because a more narrowly defined warning area (a narrow and elongated PHI plume) necessarily excludes more people who are not at risk. Some people who might otherwise ignore a text only WEA360 warning message might also be more inclined to take action in response to a message that includes a PHI forecast graphic. But other people who might take shelter in response to a dichotomous warning might instead place themselves at greater risk if their risk tolerance is high. Because warning response is a complex, complicated, and social phenomenon, the emerging PHI paradigm may therefore not necessarily produce better overall protective action decision-making for message recipients relative to NWS's current dichotomous warning paradigm. This study highlighted the myriad reasons for this situation. Message-related uncertainties and nonmessage factors are likely to erode the effectiveness of PHI-enhanced weather warning messages for some recipients. Thus, side-by-side comparison of warning messages in an operational environment—messages that include PHI forecast graphics (adapted for access and functional needs populations) and messages that are strictly dichotomous—would be warranted to assess their relative efficacy before adoption of PHI-enhanced WEA360 messages for consumers.

Acknowledgments. This project was conducted for the National Oceanic and Atmospheric Administration (NOAA), Collaborative Science, Technology, and Applied Research (CSTAR) Program, NOAA-NWS-NWSPO-2020-2006160. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NOAA. The authors are grateful to the NOAA/NWS officials, affiliated researchers, and Denver OEM officials who assisted with this project. Special thanks are given to Kyoto University's Disaster Prevention Research Institute (DPRI) faculty and students for contributing to the design of the community focus group sessions.

Data availability statement. Data from this study can be made available to researchers. Details of the data and how to request access are available from author Hamilton Bean at the University of Colorado Denver.

APPENDIX

Mock PHI-Enhanced WEA360 Messages

Examples from the set of mock messages used in the community focus groups are shown in Figs. A1–A5. The images used here were derived from Shivers-Williams et al. (2021). Table A1 lists the demographics of those community focus group participants. The study's screening questionnaire asked prospective participants to complete a shortened version of the "Tornado Knowledge Quiz" included in Casteel (2016); U.S. academic grading often uses letter grades, with A+ being the highest and F being the lowest.

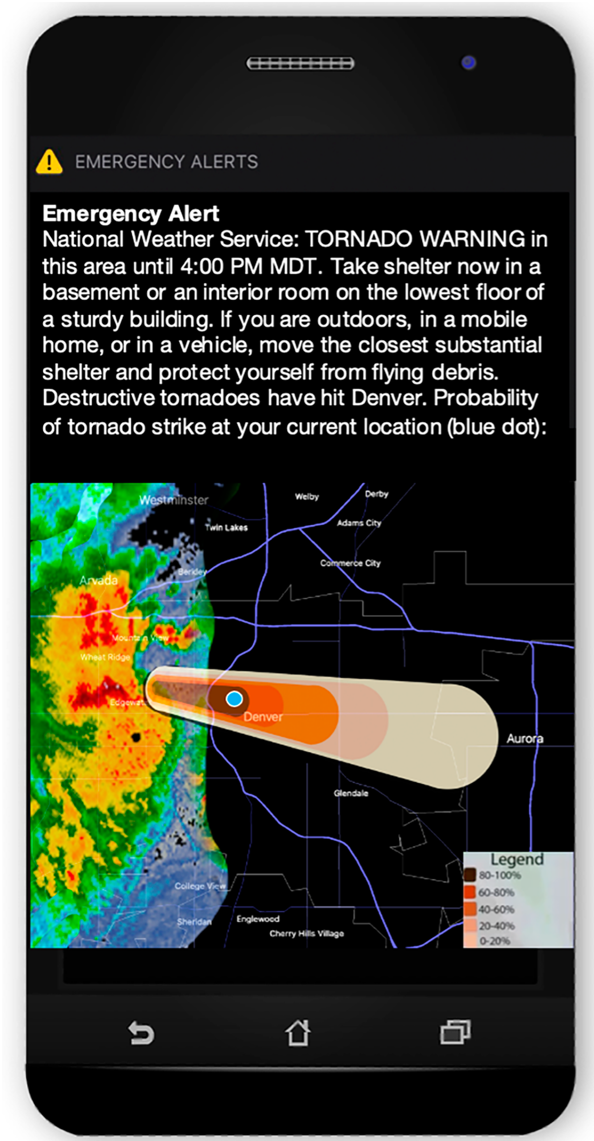

FIG. A1. Message: ~90% probability of impact.

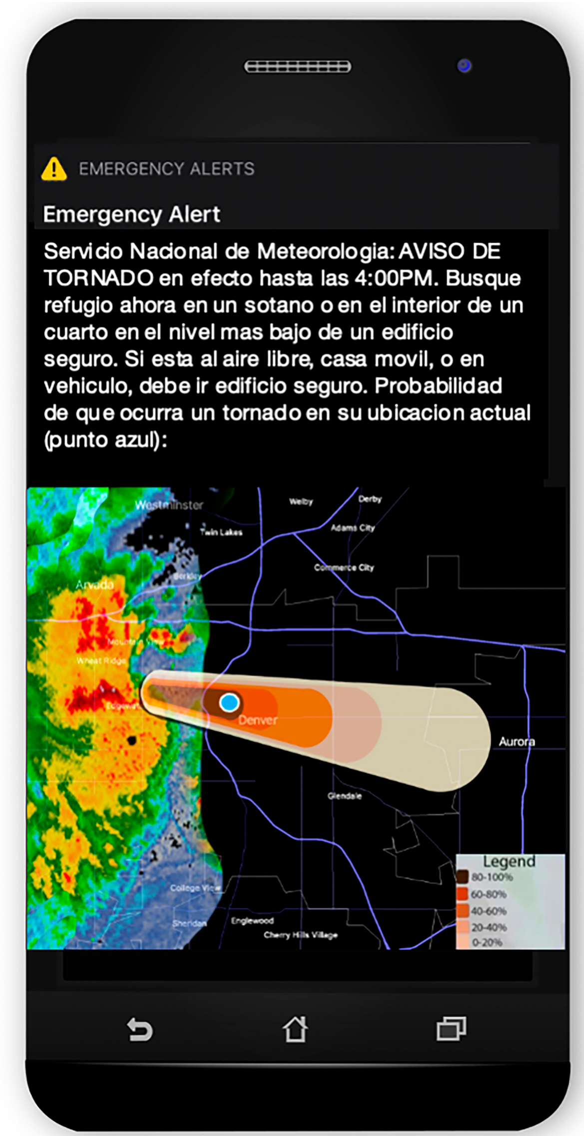

FIG. A2. Message: ~90% probability of impact, Spanish version.

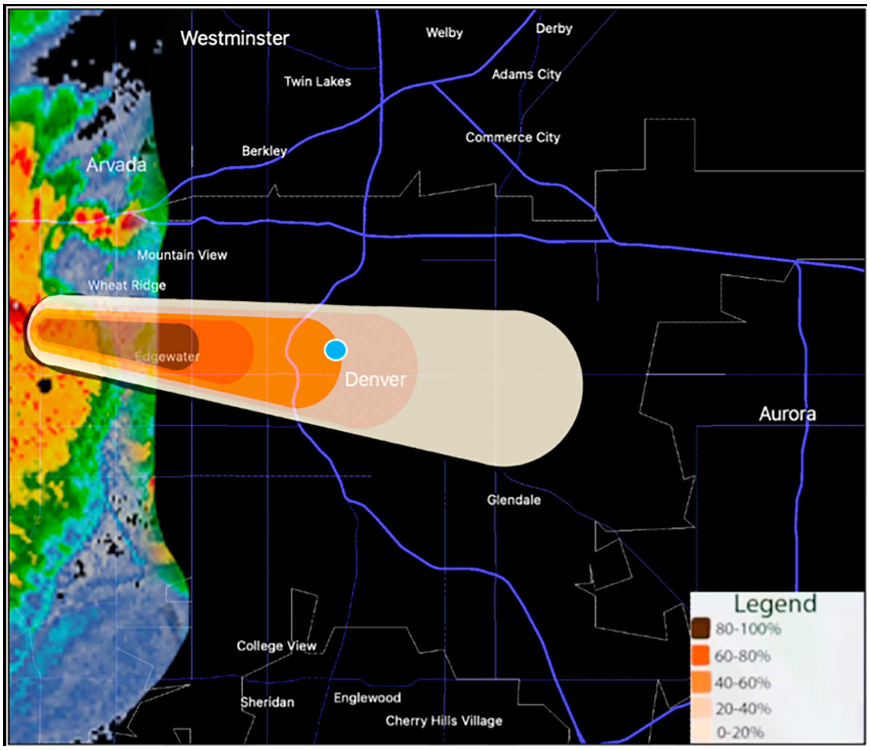

FIG. A3. ~50% probability of impact (map only—text unchanged except for time).

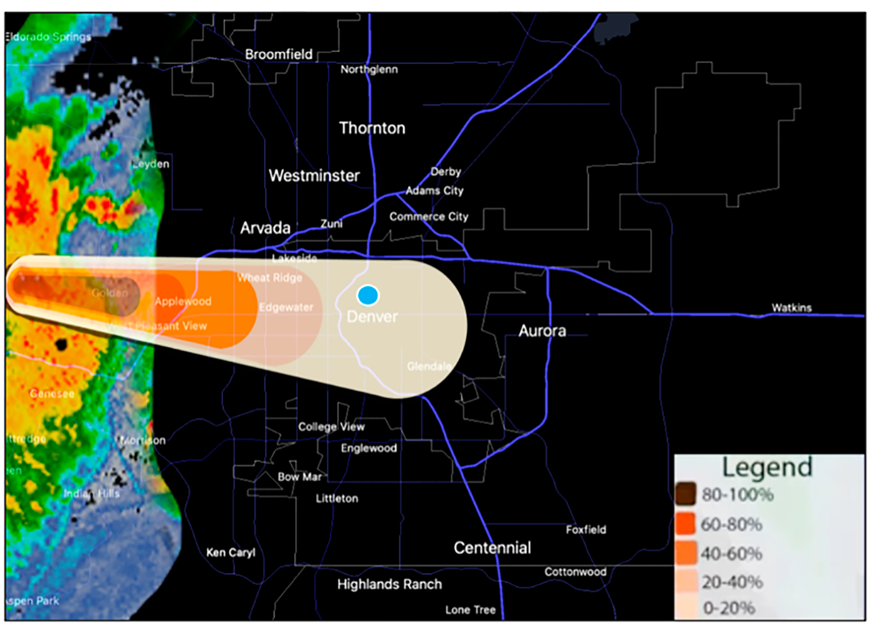

FIG. A4. ~10% probability of impact (map only—text unchanged except for time).

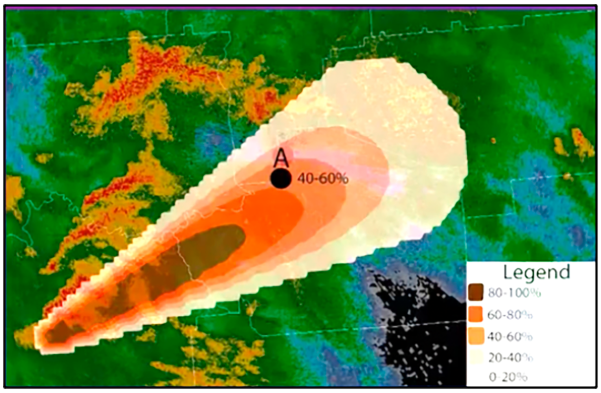

FIG. A5. PHI forecast graphic used as a model.

TABLE A1. Community focus group participant demographics.

Group	Education	Gender	Income	Age (yr)	Hispanic or Latino	Race/ethnicity	Spanish proficiency	Tornado experience	Tornado knowledge
1	Associates or technical degree	Male	$0–$18,000	26–35	No	White	Very little or none	Yes	F
1	Bachelor's degree	Female	$36,000–$59,000	26–35	No	White	High	Yes	F
1	Master's or doctorate	Male	$19,000–$35,000	61 and older	No	White	Medium	No	B
1	Associates or technical degree	Female	$0–$18,000	26–35	Prefer not to say	Prefer not to disclose	Very little or none	No	C−
2	High school/General Educational Development (GED) equivalent	Male	$36,000–$59,000	46–60	No	White	Very little or none	Yes	C−
2	Master's or doctorate	Female	$60,000–$100,000	26–35	No	White	Medium	No	A+
2	Bachelor's degree	Male	$60,000–$100,000	26–35	No	Black or African American	Very little or none	No	F
2	Bachelor's degree	Female	$0–$18,000	26–35	No	White	High	Yes	B
2	Bachelor's degree	Male	$60,000–$100,000	18–25	No	Multiethnic	Very little or none	Yes	B
3	Associates or technical degree	Female	$36,000–$59,000	26–35	Yes	Hispanic/Latino/Spanish-speaking	High	No	C−
3	Master's or doctorate	Male	$36,000–$59,000	26–35	Yes	Hispanic/Latino/Spanish-speaking	High	Yes	F
3	Associates or technical degree	Male	$36,000–$59,000	26–35	No	Multiethnic	Very little or none	Yes	C−
4	Bachelor's degree	Female	$36,000–$59,000	26–35	No	White	Low	No	F
4	No answer	Female	$0–$18,000	36–45	No	White	Very little or none	Yes	F
4	Associates or technical degree	Male	$36,000–$59,000	61 and older	Yes	Hispanic/Latino/Spanish-speaking	Low	No	C−
4	Master's or doctorate	Male	$60,000–$100,000	18–25	No	White	High	Yes	F
4	Master's or doctorate	Female	$36,000–$59,000	46–60	No	Prefer not to disclose	Medium	Yes	B
4	Associates or technical degree	Male	$36,000–$59,000	46–60	No	White	Low	Yes	B
4	Bachelor's degree	Female	$36,000–$59,000	26–35	No	White	Low	No	F
5	High school/GED equivalent	Male	$0–$18,000	26–35	No	White	Very little or none	Un.	C−
5	Bachelor's degree	Male	$36,000–$59,000	26–35	No	White	Low	No	B
5	High school/GED equivalent	Male	$0–$18,000	46–60	No	White	Low	No	C−
5	High school/GED equivalent	Male	$0–$18,000	18–25	Yes	White	Very little or none	Yes	F
5	High school/GED equivalent	Male	$19,000–$35,000	46–60	Yes	Hispanic/Latino/Spanish-speaking	Low	Yes	A+
5	Associates or technical degree	Female	$36,000–$59,000	26–35	No	White	Low	No	C−
6	High school/GED equivalent	Female	$19,000–$35,000	26–35	Yes	Hispanic/Latino/Spanish-speaking	Very little or none	No	B
6	Bachelor's degree	Female	$36,000–$59,000	26–35	No	White	Low	Yes	C−
6	Master's or doctorate	Female	$36,000–$59,000	18–25	No	White	Very little or none	No	F
6	Bachelor's degree	Male	$36,000–$59,000	36–45	No	Black or African American	Low	No	C−
6	Bachelor's degree	Male	$36,000–$59,000	26–35	No	White	Very little or none	No	B
6	Associates or technical degree	Male	$36,000–$59,000	36–45	Yes	Hispanic/Latino/Spanish-speaking	Medium	Yes	C−

REFERENCES

Aguirre, B. E., D. Wenger, and G. Vigo, 1998: A test of the emergent norm theory of collective behavior. *Sociol. Forum*, **13**, 301–320, https://doi.org/10.1023/A:1022145900928.

Bartolucci, A., C. Casareale, and J. Drury, 2021: Cooperative and competitive behaviour among passengers during the Costa Concordia disaster. *Saf. Sci.*, **134**, 105055, https://doi.org/10.1016/j.ssci.2020.105055.

Bean, H., 2019: *Mobile Technology and the Transformation of Public Alert and Warning.* Praeger, 240 pp.

——, B. Liu, S. Madden, D. Mileti, J. Sutton, and M. Wood, 2014: Comprehensive testing of imminent threat public messages for mobile devices. National Consortium for the Study of Terrorism and Responses to Terrorism, 206 pp., https://www.dhs.gov/publication/wea-comprehensive-testing-imminent-threat-public-messages-mobile-devices-updated.

——, J. Sutton, B. F. Liu, S. Madden, M. M. Wood, and D. S. Mileti, 2015: The study of mobile public warning messages: A research review and agenda. *Rev. Comm.*, **15**, 60–80, https://doi.org/10.1080/15358593.2015.1014402.

——, B. F. Liu, S. Madden, J. Sutton, M. Wood, and D. Mileti, 2016: Disaster warnings in your pocket: A qualitative study of how audiences interpret wireless emergency alerts. *J. Contingencies Crisis Manage.*, **24**, 136–147, https://doi.org/10.1111/1468-5973.12108.

——, N. Grevstad, A. Meyer, and A. Koutsoukos, 2022: Exploring whether wireless emergency alerts can help impede the spread of Covid-19. *J. Contingencies Crisis Manage.*, **30**, 185–203, https://doi.org/10.1111/1468-5973.12376.

Cain, L., E. Herovic, and K. Wombacher, 2021: "You are here": Assessing the inclusion of maps in a campus emergency alert system. *J. Contingencies Crisis Manage.*, **29**, 332–340, https://doi.org/10.1111/1468-5973.12358.

Casteel, M. A., 2016: Communicating increased risk: An empirical investigation of the National Weather Service's impact-based warnings. *Wea. Climate Soc.*, **8**, 219–232, https://doi.org/10.1175/WCAS-D-15-0044.1.

Childs, S. J., J. Demuth, R. Morss, K. Ash, S. Joslyn, C. Savelli, and C. Qin, 2021: Evaluating probabilistic tornado warning graphics: Unprimed interpretations of and responses motivated by four designs. *16th Symp. on Societal Applications: Policy, Research and Practice/Ninth Symp. on Building a Weather-Ready Nation*, Online, Amer. Meteor. Soc., 12.7, https://ams.confex.com/ams/101ANNUAL/meetingapp.cgi/Paper/381424.

Dallo, I., M. Stauffacher, and M. Marti, 2022: Actionable and understandable? Evidence-based recommendations for the design of (multi-)hazard warning messages. *Int. J. Disaster Risk Reduct.*, **74**, 102917, https://doi.org/10.1016/j.ijdrr.2022.102917.

Demuth, J. L., R. E. Morss, K. D. Ash, S. Savelli, S. Joslyn, and C. Qin, 2020: The impact of color-coded probabilistic tornado warnings on risk perceptions and responses, Part II: Interviews. *15th Symp. on Societal Applications: Policy, Research and Practice*, Boston, MA, Amer. Meteor. Soc., 3A.6, https://ams.confex.com/ams/2020Annual/meetingapp.cgi/Paper/370426.

——, J. Vickery, H. Lazrus, J. Henderson, R. E. Morss, and K. D. Ash, 2022: Rethinking warning compliance and complacency by examining how people manage risk and vulnerability during real-world tornado threats. *Bull. Amer. Meteor. Soc.*, **103**, E1553–E1572, https://doi.org/10.1175/BAMS-D-21-0072.1.

Ding, Q., and B. Millet, 2020: Visualizing uncertainty in weather forecasts. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, **64**, 1064–1068, https://doi.org/10.1177/1071181320641255.

Doermann, J. L., E. D. Kuligowski, and J. Milke, 2020: From social science research to engineering practice: Development of a short message creation tool for wildfire emergencies. *Fire Technol.*, **57**, 815–837, https://doi.org/10.1007/s10694-020-01008-7.

Drabek, T. E., and D. A. McEntire, 2003: Emergent phenomena and the sociology of disaster: Lessons, trends and opportunities from the research literature. *Disaster Prev. Manage.*, **12**, 97–112, https://doi.org/10.1108/09653560310474214.

Ericsson, K. A., and H. A. Simon, 1993: *Protocol Analysis: Verbal Reports as Data.* MIT Press, 443 pp.

FCC, 2023: Wireless Emergency Alerts. Amendments to Part 11 of the Commission's Rules Regarding the Emergency Alert System. Further Notice of Proposed Rulemaking—PS Docket Nos. 15-91 and 15-94, 83 pp., https://docs.fcc.gov/public/attachments/FCC-23-30A1.pdf.

Gillespie, K. L., D. S. LaDue, and C. D. Karstens, 2022: Tornado warning decision making with the probabilistic hazard information tool. National Weather Center Research Experience for Undergraduates, 10 pp., http://www.caps.ou.edu/reu/reu21/finalpapers/Gillespie_FinalPaper.pdf.

Jauernic, S. T., and M. S. Van Den Broeke, 2016: Perceptions of tornadoes, tornado risk, and tornado safety actions and their effects on warning response among Nebraska undergraduates. *Nat. Hazards*, **80**, 329–350, https://doi.org/10.1007/s11069-015-1970-9.

Jon, I., S. K. Huang, and M. K. Lindell, 2018: Perceptions and reactions to tornado warning polygons: Would a gradient polygon be useful? *Int. J. Disaster Risk Reduct.*, **30**, 132–144, https://doi.org/10.1016/j.ijdrr.2018.01.035.

Joslyn, S., S. Savelli, C. Qin, J. Demuth, R. Morss, and K. D. Ash, 2020: The impact of color-coded probabilistic tornado warnings on risk perceptions and responses, Part I: Experiment. *15th Symp. on Societal Applications: Policy, Research and Practice*, Boston, MA, Amer. Meteor. Soc., 3A.5, https://ams.confex.com/ams/2020Annual/meetingapp.cgi/Paper/370402.

Kim, J., A. A. Seate, B. F. Liu, D. Hawblitzel, and T. Funk, 2022: To warn or not to warn: Factors influencing National Weather Service warning meteorologists' tornado warning decisions. *Wea. Climate Soc.*, **14**, 697–708, https://doi.org/10.1175/WCAS-D-20-0115.1.

Kuligowski, E. D., and J. Doermann, 2018: A review of public response to short message alerts under imminent threat. NIST Tech. Note 1982, 39 pp., https://nvlpubs.nist.gov/nistpubs/TechnicalNotes/NIST.TN.1982.pdf.

LaForce, S., and D. Bright, 2021a: Are we there yet? The developing state of mobile access equity. *Assistive Technol.*, **31**, 628–636, https://doi.org/10.1080/10400435.2021.1926372.

——, and ——, 2021b: Evaluating the impact of WEA 2.0 regulations on WEA message content accessibility. *J. Emerg. Manage. Disaster Commun.*, **2**, 239–257, https://doi.org/10.1142/S2689980921500135.

Lambropoulos, D., M. Yousefvand, and N. Mandayam, 2021: Tale of seven alerts: Enhancing Wireless Emergency Alerts (WEAs) to reduce cellular network usage during disasters. arXiv, 2102.00589v1, https://doi.org/10.48550/arXiv.2102.00589.

Lee, M., and M. You, 2021: Effects of COVID-19 emergency alert text messages on practicing preventive behaviors: Cross-sectional web-based survey in South Korea. *J. Med. Internet Res.*, **23**, e24165, https://doi.org/10.2196/24165.

Lindell, M. K., and T. C. Earle, 1983: How close is close enough: Public perceptions of the risks of industrial facilities. *Risk Anal.*, **3**, 245–253, https://doi.org/10.1111/j.1539-6924.1983.tb01393.x.

——, and R. W. Perry, 2012: The protective action decision model: Theoretical modifications and additional evidence. *Risk Anal.*, **32**, 616–632, https://doi.org/10.1111/j.1539-6924.2011.01647.x.

——, S. Arlikatti, and S. K. Huang, 2019: Immediate behavioral response to the June 17, 2013 flash floods in Uttarakhand, North India. *Int. J. Disaster Risk Reduct.*, **34**, 129–146, https://doi.org/10.1016/j.ijdrr.2018.11.011.

——, C. Chen, and M. J. Shaw, 2022: Modeling hurricane evacuation warnings: Effects of message content and timing on risk perception and response. *National Hurricane Conf.*, Orlando, FL, National Hurricane Conference, https://www.hurricanemeeting.com/download/Hurricane-Hazard-Perception-and-Risk-Communications-4.12.2022.pdf.

Ling, R., and B. Oppegaard, 2021: THIS IS NOT A DRILL: Mobile telephony, information verification, and expressive communication during Hawaii's false missile alert. *Soc. Media Soc.*, **7** (1), 1–12, https://doi.org/10.1177/2056305121999661

Liu, B. F., M. M. Wood, M. Egnoto, H. Bean, J. Sutton, D. Mileti, and S. Madden, 2017: Is a picture worth a thousand words? The effects of maps and warning messages on how publics respond to disaster information. *Public Relat. Rev.*, **43**, 493–506, https://doi.org/10.1016/j.pubrev.2017.04.004.

McBride, S. K., and Coauthors, 2021: Evidence-based guidelines for protective actions and earthquake early warning systems. *Geophysics*, **87**, 1–79, https://doi.org/10.1190/geo2021-0222.1.

Mileti, D. S., 2018: PrepTalks: Modernizing public warning messaging. FEMA, https://www.fema.gov/media-library/assets/videos/159069.

——, and J. H. Sorensen, 1990: Communication of emergency public warnings: A social science perspective and state-of-the-art assessment. Oak Ridge National Laboratory Rep. ORNL-6609, 162 pp., https://www.osti.gov/biblio/6137387-communication-emergency-public-warnings-social-science-perspective-state-art-assessment.

——, and L. Peek, 2000: The social psychology of public response to warnings of a nuclear power plant accident. *J. Hazard. Mater.*, **75**, 181–194, https://doi.org/10.1016/S0304-3894(00)00179-5.

Miran, S. M., C. Ling, J. J. James, A. Gerard, and L. Rothfusz, 2017a: User perception and interpretation of tornado probabilistic hazard information: Comparison of four graphical designs. *Appl. Ergon.*, **65**, 277–285, https://doi.org/10.1016/j.apergo.2017.06.016.

——, ——, ——, ——, and ——, 2017b: Effective method to convey threat information for tornado: Probabilistic hazard information vs. deterministic hazard information. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, **61**, 292–296, https://doi.org/10.1177/1541931213601554.

——, ——, A. Gerard, and L. Rothfusz, 2018: The effect of providing probabilistic information about a tornado threat on people's protective actions. *Nat. Hazards*, **94**, 743–758, https://doi.org/10.1007/s11069-018-3418-5.

——, ——, ——, and ——, 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 Anal.*, **39**, 1533–1545, https://doi.org/10.1111/risa.13289.

——, ——, and J. J. James, 2020: People's thresholds of decision-making against a tornado threat using dynamic probabilistic hazard information. *Int. J. Disaster Risk Reduct.*, **42**, 101345, https://doi.org/10.1016/j.ijdrr.2019.101345.

——, ——, and S. Datta, 2021: Bayesian hierarchical modeling of people's decision-making during an extreme weather event. *Jpn. J. Stat. Data Sci.*, **4**, 411–425, https://doi.org/10.1007/s42081-021-00114-2.

Morss, R. E., C. L. Cuite, J. L. Demuth, W. K. Hallman, and R. L. Shwom, 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. *Int. J. Disaster Risk Reduct.*, **30**, 44–58, https://doi.org/10.1016/j.ijdrr.2018.01.023.

National Academies of Sciences, Engineering, and Medicine, 2018: *Emergency Alert and Warning Systems: Current Knowledge and Future Research Directions*. National Academies Press, 128 pp., https://www.nap.edu/catalog/24935/emergency-alert-and-warning-systems-current-knowledge-and-future-research.

National Severe Storms Laboratory, 2020: FACETs: Forecasting a Continuum of Environmental Threats. NOAA, https://www.nssl.noaa.gov/projects/facets/.

Obermeier, H. B., K. L. Berry, K. E. Klockow-McClain, A. Campbell, C. Carithers, A. Gerard, and J. E. Trujillo-Falcón, 2022: The creation of a research television studio to test probabilistic hazard information with broadcast meteorologists in NOAA's Hazardous Weather Testbed. *Wea. Climate Soc.*, **14**, 949–963, https://doi.org/10.1175/WCAS-D-21-0171.1.

O'Connor, C., and H. Joffe, 2020: Intercoder reliability in qualitative research: Debates and practical guidelines. *Int. J. Qual. Methods*, **19**, 1–19, https://doi.org/10.1177/1609406919899220.

Phillips, N., and C. Hardy, 2002: *Discourse Analysis: Investigating Processes of Social Construction.* Sage, 97 pp.

Ripberger, J. T., C. L. Silva, H. C. Jenkins-Smith, J. Allan, M. Krocak, W. Wehde, and S. Ernst, 2020: Exploring community differences in tornado warning reception, comprehension, and response across the United States. *Bull. Amer. Meteor. Soc.*, **101**, E936–E948, https://doi.org/10.1175/BAMS-D-19-0064.1.

Sellnow, T. L., and M. W. Seeger, 2021: *Theorizing Crisis Communication.* John Wiley and Sons, 368 pp.

Shivers-Williams, C. A., and K. E. Klockow-McClain, 2020: Geographic scale and probabilistic forecasts: A trade-off for protective decisions? *Nat. Hazards*, **105**, 2283–2306, https://doi.org/10.1007/s11069-020-04400-2.

——, K. E. Klockow, and K. Berry, 2021: Communicating uncertainty: A public perspective on probabilistic forecast products. *16th Symp. on Societal Applications: Policy, Research and Practice*, Online, Amer. Meteor. Soc., 1.4, https://ams.confex.com/ams/101ANNUAL/meetingapp.cgi/Paper/378869.

Smith, K. R., S. Grant, and R. E. Thomas, 2022: Testing the public's response to receiving severe flood warnings using simulated cell broadcast. *Nat. Hazards*, **112**, 1611–1631, https://doi.org/10.1007/s11069-022-05241-x.

Smith, L., and T. Horner, 2021: Tornadoes in Denver: The Denver convergence-vorticity zone (DCVZ) and Denver cyclone. Highpoint Weather, https://medium.com/14ers-forecast/tornadoes-in-denver-the-denver-convergence-vorticity-zone-dcvz-and-denver-cyclone-c57dc0468df6.

Strauss, A., and J. M. Corbin, 1997: *Grounded Theory in Practice.* Sage, 288 pp.

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

Sutton, J., and C. Woods, 2016: Tsunami warning message interpretation and sense making: Focus group insights. *Wea. Climate Soc.*, **8**, 389–398, https://doi.org/10.1175/WCAS-D-15-0067.1.

——, and E. D. Kuligowski, 2019: Alerts and warnings on short messaging channels: Guidance from an expert panel process.

Nat. Hazards Rev., **20**, 2, https://doi.org/10.1061/(ASCE)NH.1527-6996.0000324.

——, and L. M. Fischer, 2020: Understanding visual risk communication messages: An analysis of visual attention allocation and think aloud responses to tornado graphics. *Wea. Climate Soc.*, **13**, 173–188, https://doi.org/10.1175/WCAS-D-20-0042.1.

——, ——, and M. M. Wood, 2021: Tornado warning guidance and graphics: Implications of the inclusion of protective action information on perceptions and efficacy. *Wea. Climate Soc.*, **13**, 1003–1014, https://doi.org/10.1175/WCAS-D-21-0097.1.

Takenouchi, K., C. Nakanishi, K. Yamori, M. Sawada, K. Takeuchi, and H. Fujiwara, 2017: Collaborative Community Weather Information for meteorological disasters: A case study of Nakajima SchoFol District, Ise. *J. Intergr. Disaster Risk Manage.*, **7** (2), 1–24, https://doi.org/10.5595/idrim.2017.0200.

Tracy, S. J., 2019: *Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact.* John Wiley and Sons, 368 pp.

Trujillo-Falcón, J. E., O. Bermúdez, K. Negrón-Hernández, J. Lipski, E. Leitman, and K. Berry, 2021: Hazardous weather communication en Español: Challenges, current resources, and future practices. *Bull. Amer. Meteor. Soc.*, **102**, E765–E773, https://doi.org/10.1175/BAMS-D-20-0249.1.

——, J. Reedy, K. E. Klockow-McClain, K. L. Berry, G. J. Stumpf, A. V. Bates, and J. G. LaDue, 2022: Creating a communication framework for FACETs: How probabilistic hazard information affected warning operations in NOAA's hazardous weather testbed. *Wea. Climate Soc.*, **14**, 881–892, https://doi.org/10.1175/WCAS-D-21-0136.1.

Walters, J. E., L. R. Mason, and K. N. Ellis, 2019: Examining patterns of intended response to tornado warnings among residents of Tennessee, United States, through a latent class analysis approach. *Int. J. Disaster Risk Reduct.*, **34**, 375–386, https://doi.org/10.1016/j.ijdrr.2018.12.007.

——, ——, K. Ellis, and B. Winchester, 2020: Staying safe in a tornado: A qualitative inquiry into public knowledge, access, and response to tornado warnings. *Wea. Forecasting*, **35**, 67–81, https://doi.org/10.1175/WAF-D-19-0090.1.

Wehde, W., J. T. Ripberger, H. Jenkins-Smith, B. A. Jones, J. N. Allan, and C. L. Silva, 2021: Public willingness to pay for continuous and probabilistic hazard information. *Nat. Hazards Rev.*, **22**, 04021004, https://doi.org/10.1061/(ASCE)NH.1527-6996.0000444.

Wood, M. M., D. S. Mileti, H. Bean, B. F. Liu, J. Sutton, and S. Madden, 2018: Milling and public warnings. *Environ. Behav.*, **50**, 535–566, https://doi.org/10.1177/0013916517709561.