Understanding Broadcast Meteorologists' Current and Future Use of Severe Weather Watches, Warnings, and Probabilistic Hazard Information

HOLLY B. OBERMEIER,[a,b] KODI L. BERRY,[b] AND JOSEPH E. TRUJILLO-FALCÓN[a,b]

[a] *Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma, Norman, Oklahoma*
[b] *NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma*

(Manuscript received 7 February 2023, in final form 7 August 2023, accepted 14 August 2023)

ABSTRACT: Broadcast meteorologists are essential in the communication of National Weather Service (NWS) warnings to the public. Therefore, it is imperative to include them in a user-centered approach for the design and implementation of new warning products. Forecasting a Continuum of Environmental Threats (FACETs) will modernize the way meteorologists forecast and communicate NWS warning information to the general public using rapidly updating probabilistic hazard information (PHI). Storm-scale PHI consists of probabilistic forecasts for severe wind/hail, tornadoes, and lightning hazards. Hence, NWS warnings would have the capacity to be supplemented by a quantitative or qualitative likelihood of hazard occurrence. The researchers conducting this study wanted to know what broadcast meteorologists thought about the inclusion of this likelihood information and how it could impact their decision-making and communication process. Using a nationwide survey, this team of researchers first asked broadcast meteorologists about their current practices for severe weather coverage using NWS watches and warnings. Next, broadcast meteorologists were introduced to multiple iterations of PHI prototypes and queried for their input. Findings indicated that broadcast meteorologists already face a complex decision-making and communication process under today's warning paradigm. In addition, respondents were split on whether to explicitly communicate probabilities with their viewers. Respondents' choices were also somewhat inconclusive regarding nomenclature, definitions of PHI and representations of PHI with warning polygons. These results suggest that PHI should feature user-driven, customizable options to fulfill broadcast meteorologists' needs and that the iterative nature of the research-and-development process of PHI should continue.

SIGNIFICANCE STATEMENT: Broadcast meteorologists are vital communicators of dangerous weather to the public, leading researchers to study them more closely. Using a nationwide survey, this team of researchers wanted to know how broadcast meteorologists talk about tornadoes, large hail, and high winds to their viewers under today's system of National Weather Service warnings. Survey findings indicated that broadcast meteorologists face a complex decision-making process when communicating dangerous weather. Any effort to modernize the current warning system, such as including hazard probability, should consider this complex process. Modernization should complement the role of broadcast meteorologists to ultimately serve the public and user-driven options should be a key component of any probabilistic information that is included in a future National Weather Service warning system.

KEYWORDS: Severe storms; Probability forecasts/models/distribution; Broadcasting; Communications/decision-making; Societal impacts

1. Introduction and background

Television broadcasts continue to be a leading source of National Weather Service (NWS) tornado warning information for the public (Silva et al. 2017, 2018, 2019; Krocak et al. 2020, 2021; Bitterman et al. 2022). Despite shifting audience preferences in the age of mobile devices, broadcast meteorologists remain relevant and maintain a presence on a variety of platforms. Beyond traditional on-air coverage, broadcasters are responsible for updating audiences on severe weather threats on social media, mobile applications, website channels, and radio. These important intermediaries provide and explain severe weather information to their viewers and are considered important partners of the NWS. Knowing the weather information needs of this NWS core partner group is paramount in any effort to understand, update or change the existing severe weather alerting timeline of outlooks, watches, and warnings. Forecasting a Continuum of Environmental Threats (FACETs; Rothfusz et al. 2018) proposes the introduction of more rapidly updating warnings, supplemented with probabilistic hazard information (PHI; Karstens et al. 2014, 2015, 2018). PHI is defined as "the probability for a given hazardous weather phenomena within a defined spatial and temporal range" (Karstens et al. 2014). As part of efforts to ensure that PHI is implemented in an effective way, researchers have prioritized studying broadcast meteorologists, especially within NOAA's Hazardous Weather Testbed (Nemunaitis-Berry and Obermeier 2017; Kolakowski et al. 2019; Calhoun et al. 2021; Obermeier et al. 2017, 2018, 2020, 2022). Within these Hazardous Weather Testbed studies, storm-based probabilistic

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/WCAS-D-23-0013.s1.

Corresponding author: Holly Obermeier, holly.obermeier@ou.edu

DOI: 10.1175/WCAS-D-23-0013.1

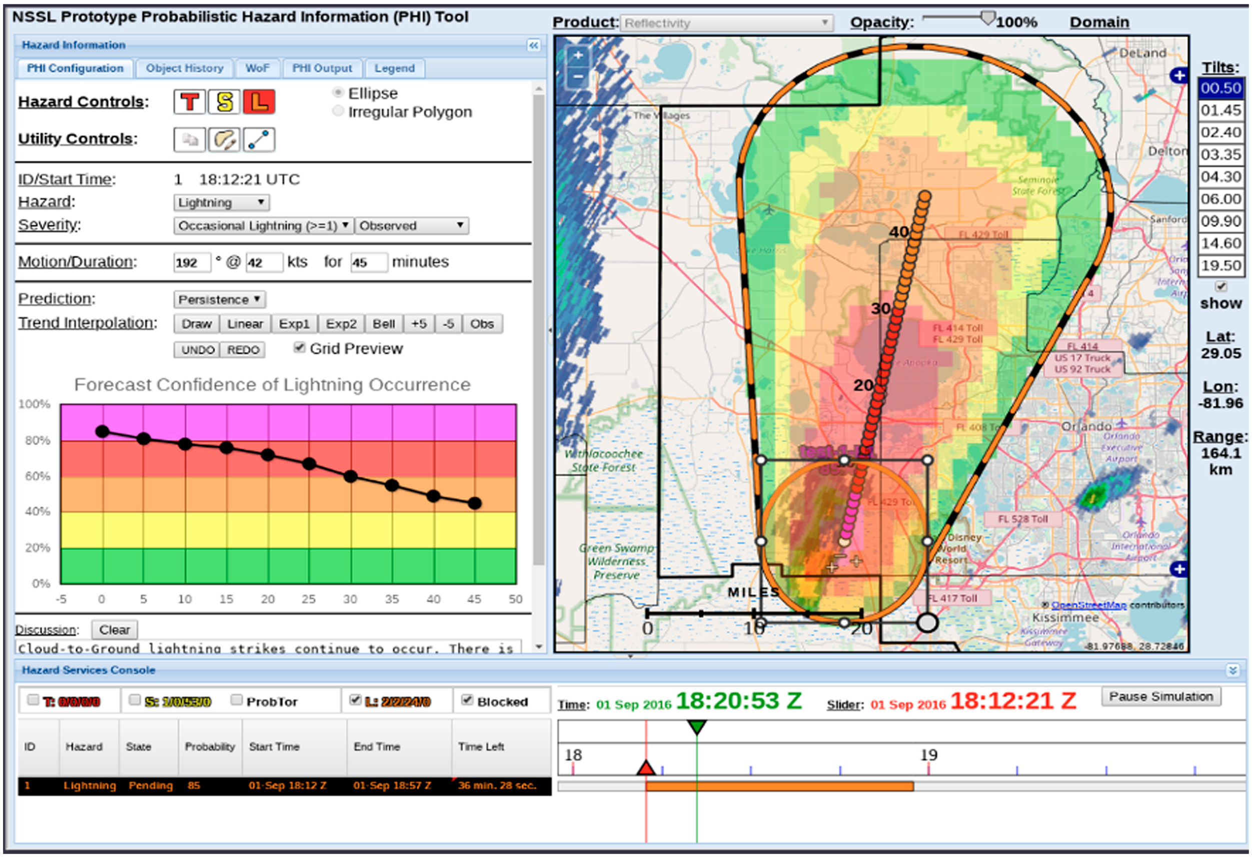

FIG. 1. A display of the prototype tool in which PHI can be created by an NWS forecaster. The forecaster can manipulate the storm vector and a trend line of probability over time (left side of display) as well as the shape and scale of the PHI (right side of the display).

forecasts for severe wind/hail, tornadoes, and lightning hazards (known as ProbSevere for wind and hail; Cintineo et al. 2014; Meyer et al. 2019; Calhoun et al. 2018) were created and/or modified by NWS forecasters to produce grid-based probabilistic forecasts (or PHI) within a prototype software tool (Fig. 1). PHI moved with the hazards in space, updating through either machine automation and/or forecaster modification every 2 min. PHI was then displayed to broadcast meteorologists, who used the information to make decisions in a simulated television studio environment. These types of decisions included running crawling messages, providing short on-air cut-ins (generally <5 min), providing continuous on-air coverage ("wall to wall") and making social media posts (Obermeier et al. 2022). Findings indicated that broadcast meteorologists found PHI to be valuable, in that PHI provided more specific hazard and timing information than warning polygons alone (Obermeier et al. 2022). According to participants, PHI helped supplement the information voids that can occur during the warning scale, and gave them more contextualized information to share with their viewers rather than relying on the binary "yes or no" nature of warning polygons (Obermeier et al. 2022).

Testing of PHI within the Hazardous Weather Testbed represented a rapid prototyping environment, in which quick changes could be made to PHI in situ (Karstens et al. 2014). The advantage was that PHI was situated in a naturalistic environment representing the workplace and many of the complexities of the real world, strengthening the usability aspect of these studies (Karstens et al. 2018). Quick changes could be made to the prototype during and between experiment cycles. Through observations and assessments from broadcast meteorologists participating in the Hazardous Weather Testbed, researchers have identified several areas in need of additional research and iterative development in regard to PHI. These areas include the definition of PHI, probabilistic thresholds for decision-making, the nomenclature of PHI, update frequency, qualitative and quantitative representations of PHI, and the relationship of PHI with current severe weather warnings.

Concerning the definition of PHI, a few iterations have been used within the studies mentioned above in the Hazardous Weather Testbed. One of these definitions included the "probability of a warning being issued" (e.g., *The National Weather Service projects a 75% probability of issuing a tornado warning within the next hour*). This definition evolved from both local NWS probabilistic thunderstorm grids and the development of the Warn-on-Forecast System (Hatfield et al. 2018, 2023; Wilson et al. 2023). A second definition of

PHI was the "probability of hazard" (e.g., *The National Weather Service projects a 75% probability of a tornado for a location within the next hour*). This was the developer definition mentioned earlier in the introduction as "the probability for a given hazardous weather phenomena within a defined spatial and temporal range" (Karstens et al. 2014). A third definition of PHI was "confidence in hazard occurrence" (e.g., *a National Weather Service forecaster is 75% confident in this storm producing a tornado within the next hour*). During the 2017 PHI prototype experiment within the Hazardous Weather Testbed, NWS forecasters expressed the need to convey both hazard probability and confidence in a singular way to emergency managers. Therefore, hazard probabilities were framed as "forecast confidence of hazard occurrence" in that experiment, which appeared to effectively align with both forecasters' and emergency managers' needs (Karstens et al. 2018). Broadcast meteorologists have not yet been queried more specifically about this definition. Alternatively, it could be possible that an entirely different description of PHI may be more useful for the broadcast meteorologist community.

Building off a general definition of PHI, an additional suggestion of the FACETs framework is the establishment of specific thresholds of PHI. Because of the number of temporal and spatial reference classes, a seemingly infinite number of options could be possible (Rothfusz et al. 2018). In previous usability research in the Hazardous Weather Testbed, some broadcast participants have voiced a desire for more specificity in probabilistic products based upon the Storm Predictions Center's definition of significant severe hazards [wind speeds of 65 kt (1 kt ≈ 0.51 m s^{-1}) or greater; hail 2 in. in diameter or larger; Obermeier et al. 2022]. Beyond ascertaining PHI definitions and thresholds, pinpointing a more usable term for PHI has also been a subject of testing. Developers have commonly referred to PHI as a "plume" or "swath" of probabilities (Karstens et al. 2018). Broadcast meteorologists have told researchers that these terms are jargon, and that they would not use these terms when communicating with their viewers (Obermeier et al. 2022). When testing PHI in the 2019 Hazardous Weather Testbed, researchers began intentionally tracking the different kinds of words broadcasters were using to describe the plume during their mock on-air severe weather coverage (Obermeier et al. 2020, 2022). A wide range of descriptors emerged (some more formal sounding than others), including "cone," "egg," "bubble," "bullseye," "layer cake," "big yellow blob," "cone of certainty," and "cone of danger." According to one Spanish-speaking participant, the word "plume" could not be translated from English in an understandable way. Therefore, this individual used the terms "círculo de peligro" and "cono de posibilidad" (in English, "danger circle" and "cone of possibility," respectively). Nomenclature matters, because how a broadcast meteorologist names PHI could impact how their audience perceives and understands the product and the "weather story" the broadcaster is trying to communicate (Obermeier et al. 2022; Trujillo-Falcón et al. 2022a).

In today's warning paradigm, an average of 20.8 min pass from the time of warning issuance to the first update to that warning (Harrison and Karstens 2017). Developers designed storm-scale PHI to automatically update as quickly as every 2 min through the inclusion of algorithmic automation (Karstens et al. 2018). Hence, PHI represents a tenfold increase in update frequency for severe convective hazards information at that scale. While participants in many Hazardous Weather Testbed experiments wanted to receive updates as swiftly as possible, the idea of information arriving so frequently came with concerns. Within these experiments, broadcast meteorologist participants often became overwhelmed and mentioned the PHI was "moving and wobbling" too much at the 2-min frequency (Obermeier et al. 2022). In addition, broadcasters struggled to keep crawls and social media up to date with the most recent information (Obermeier et al. 2022). Over successive years of testing PHI in the Hazardous Weather Testbed, test cases evolved to give broadcast participants the option to choose their preferred update frequency (2, 5, 10, or 20 min). Often the 2-min option was selected, but participants noted that such frequent updates were not always necessary for severe hazards such as wind and hail and that 5- or 10-min updates could be sufficient (Obermeier et al. 2022). On the other hand, during tornado cases, broadcasters appreciated having the quick flow of information to share with their viewers during wall to wall coverage and to justify staying on the air to their management (Obermeier et al. 2022).

Early testing of PHI with broadcast meteorologists revealed the need for a qualitative labeling option (Nemunaitis-Berry et al. 2017). Most frequently, "low," "medium," and "high" labels were requested. While this type of labeling represents a simple Likert scale, research with members of the public indicates that using rank adjectives (such as "low," "medium," "high") can be an effective way to indicate the magnitude of probability (Lenhardt et al. 2020). Per software ability, subsequent testing included this quantitative representation of probability in addition to percentages (Obermeier et al. 2022). In the testbed, participants could choose between the two to share with their viewers. Ultimately, many of these participants were wary of sharing probabilities for severe weather hazards as percentages with their viewers. While they themselves wanted to see the numeric PHI, they would often share the qualitative formats of PHI on their coverage platforms. Other studies have shown that a mix of both nonnumeric and numeric probabilities is likely the most effective way to communicate to nonheterogeneous audiences (Budescu et al. 2014; Kox et al. 2015; Lenhardt et al. 2020; Ripberger et al. 2022).

Beyond PHI labels, visual representation challenges have also emerged during prior testbed studies. Storm-scale PHI is designed to supplement official NWS watch and warning products. How these products ultimately appear with one another is still under question, and how the two products are displayed could have an eventual impact on interpretation. During usability testing in the Hazardous Weather Testbed, broadcasters most frequently showed the two products layered with one another when displaying on the air (Obermeier et al. 2022). Previous research on warning polygons indicates that people perceive different levels of risk near and within the boundaries of the polygon and often experience a subsequent centroid effect (Lindell 2020). In addition, the shape of the polygon (specifically length) can have an impact on perception (Klockow-McClain et al. 2020). While PHI is meant to help bring context to the polygon, it is unknown whether PHI will mitigate or enhance legacy polygon effects, especially for a television viewing audience. Researchers have used the Testbed environment to experiment with multiple iterations of polygon and plume. These included drawing a warning polygon from a specific PHI threshold (e.g., creating a polygon from the 20% outline) and drawing a warning polygon from the outer boundary of PHI (Karstens et al. 2015, 2018). While testing PHI in the 2018 Hazardous Weather Testbed, researchers gave broadcast meteorologists both the plume and the actual warning polygon issued by the NWS in displaced real-time cases (Obermeier et al. 2019). During the 2019 Hazardous Weather Testbed experiment, forecasters could issue warning polygons derived from the PHI, in which the warning could extend beyond the PHI plume or the PHI plume could extend beyond the warning (Obermeier et al. 2020). Each of these methods presented unique challenges; in particular, broadcast meteorologists struggled to explain why some probability values were included in the warning polygon for certain warnings, but not for others. Yet, broadcasters felt that when the PHI plume extended beyond the warning or when PHI trends were increasing, they could give certain viewers extra lead time to seek shelter (Obermeier et al. 2022). Additionally, the real-world setting of the testbed could result in all sorts of visual configurations of the polygon and the PHI plume, some of which could make drastic changes in shape due to the 2-min updates and due to automation and forecaster intervention.

Engaging NWS core partners well before the transition to operations of any new product, such as PHI, is a key component in widespread product adoption and success (Robertson and Droegemeier 1990). Usability studies with many NWS core partners in the testbed environment have been comprehensive over many years, and results have been integrated into product development cycles (Nemunaitis-Berry et al. 2020). These cycles occur before the transition to operations, to avoid unintended outcomes that may disadvantage the user. Therefore, usability studies, such as the PHI prototype experiments in the Hazardous Weather Testbed described above, are critically important. However, these studies are somewhat limited by small sample sizes due to logistic and funding constraints. To explore the testbed findings on a broader scale, researchers created the survey described in this paper to deploy to broadcast meteorologists nationwide. Another goal of the survey was to confirm how broadcast meteorologists currently utilize NWS watch, warning, and outlook products in their workflow. This team of researchers had previously explored the daily structure of broadcasters' workflow during severe weather events using cognitive task analysis (CTA; Ernst 2020). The CTAs were conducted with 15 broadcast meteorologists according to the Storm Prediction Center's day one convective outlook levels, including no risk, slight risk, and moderate/high risk. On moderate and high risk days, these broadcasters reported that they often experienced a good deal of prework anxiety and would typically arrive to work up to 2 h before their scheduled shifts (Ernst 2020). To meet the demands of severe weather coverage, these broadcast meteorologists typically increased their staffing by nearly two people on moderate and high risk days, reporting that they must juggle forecast information and communicating with their newsrooms and viewers (Ernst 2020). It may be that the addition of storm-scale PHI will have impacts on broadcasters' workflow and staffing; therefore, gaining additional understanding of their duties and the technology broadcasters currently operate within is needed. The survey presented in this paper also addressed topics such as limitations with crawl systems, closed captioning abilities, and coverage decisions and strategies. In addition, this team of researchers posed questions about the types of NWS products broadcasters preferred to share most frequently and how they chose to share them. Data gathered in this survey corroborate data collected in testbed studies to help confirm exploratory findings and indicate areas that require further iterative research.

2. Method and data

Researchers created a 53-question survey focused on severe convective weather and PHI. The survey contained three primary sections, including 1) current practices for coverage of severe convective weather and Spanish translations of current NWS products, 2) PHI and its relation to current NWS severe convective warnings, and 3) demographic/background information. Completion time averaged 42 min. Question types included randomized multiple choice response lists, matrices, and open-ended questions. Respondents could choose multiple answers to many of the questions, and none were forced answers. Questions regarding current practices addressed crawling messages, weather radios, closed captioning, coverage decisions, social media, display colors and language. Respondents were shown prototypes of PHI that mirrored those used in the Hazardous Weather Testbed and were asked specific questions regarding the definition of PHI, probabilistic thresholds for decision-making, the nomenclature of PHI and update frequency. Respondents were also shown quantitative and qualitative representations of PHI. Quantitative probabilities were indicated by percentages, while qualitative probabilities were indicated by a simple three point Likert scale; both were similar to prototypes tested in the Hazardous Weather Testbed. Respondents were also shown prototypes of PHI in tandem with different abstractions of warning polygons. The scales and combinations of the two products were based on examples observed within the 2018 Hazardous Weather Testbed, in which warning polygons and PHI plumes would vary in size and position in relation to one another (Obermeier et al. 2019). These examples were designed to gain first impressions of the two products shown simultaneously. The shape of the PHI plume itself was kept constant. Survey questions are provided in the online supplemental material.

Researchers designed and developed the online survey using Qualtrics software. A survey pretest was conducted with two broadcast professionals to check for any errors and confirm clarity. This study was approved under the University of Oklahoma's Institutional Review Board (IRB 13320). Survey distribution began in December 2020. Links were sent to an email list of broadcast meteorologists using the University of Oklahoma's subscription to the nationwide Cision media database (1307 sent; 218 undeliverable; http://www.cision.com/us/pr-software/media-database). The Cision media database includes contact information for every broadcast meteorologist currently employed across the United States, and therefore should represent the entire population. In addition, the research team shared links with the National Weather Association seal holders, a Facebook group of over 2000 broadcast meteorologists, cooperative institute Twitter accounts, the American Meteorological Society (AMS) Women's Slack channel, and the Advanced Warning and Response Network (AWARN) Alliance (AWARN 2022). To gain better access to U.S. Spanish speakers, the researchers also forwarded the survey to the AMS Committee for Hispanic and Latinx Advancement through Slack and email (Trujillo-Falcón et al. 2021).

A total of 168 survey responses were collected; 115 of these were complete responses. Fifty-three (53) were not fully completed, but the answered questions in these partial surveys still provided valuable data. Responses came from all 50 states, with an estimated response rate of less than 10%. The exact response rate is difficult to quantify because of the variety of distribution channels, however, other surveys of this population have had similar rates (Reed and Senkbeil 2020; Perkins et al. 2020). Researchers also expected low participation because of the survey timing during the COVID-19 pandemic, under which members of the broadcast media were facing several abnormal constraints in their working environments. In addition, because of email filtering for some corporate entities, it is possible the survey email may have been marked as spam and therefore not received.

Table 1 outlines the demographics of survey participants. Of the total responses, 75% (n = 84) were male, 22% (n = 25) were female, and 3% (n = 3) preferred not to answer. Most respondents (32%) were 30–39 years of age, although the total age range included those from 18 to 70 years of age. Respondents could choose their race and ethnicity from a list (see the online supplemental material). The racial distribution of respondents included Black or African American (3%, n = 3), White (96%; n = 105), and other races (3%; n = 2). Hispanic and/or Latinx respondents represented 4% (n = 4) of the total. Reported job titles of respondents reflected the myriad of shifts for which broadcast meteorologists are responsible. While 37% (n = 41) held the position of chief meteorologist, many others held meteorologist positions with responsibilities for morning and weekend shifts (some of which included additional general reporting duties). Some respondents (34%; n = 38) chose not to indicate their shift, simply choosing the title of meteorologist. Sixty-five percent (65%; n = 55) of respondents had over 10 years of experience working in their field. Of those who participated, 59% (n = 67) held an AMS Seal of Approval, 29% (n = 32) held a National Weather Association (NWA) Seal of Approval, and 48% (n = 54) were AMS Certified Broadcast Meteorologists (note that some participants held multiple certificates). It is difficult to ascertain if the demographics of this survey are truly representative of the broadcast meteorologist population, as few studies of this group exist. Cranford (2018) indicates that women comprised 29% of all weathercaster positions in 2018, meaning women were somewhat underrepresented in this study. In several nationwide surveys regarding broadcast meteorologists' views on climate change (Perkins et al. 2020), the age and job title demographics of respondents were like those in this study.

Upon completion of the survey, researchers derived descriptive statistics from the data through both Qualtrics and Microsoft Excel. These statistics were drawn from the total number of respondents for the particular question. Many of the survey questions resulted in answers that were qualitative in nature. To analyze these qualitative data, a single coder conducted a thematic analysis of the responses to inductively create codes and draw out common themes (Braun and Clarke 2006). Specific examples from respondents were used to illustrate these themes.

TABLE 1. Demographics of survey respondents.

Category	Percentage of respondents
Age	
18–29	23%
30–39	32%
40–49	17%
50–59	15%
60–69	13%
70+	1%
Gender	
Male	75%
Female	22%
Other, please specify	3%
Race	
Black or African American	3%
White	95%
Other	2%
Job title	
Meteorologist	34%
Chief meteorologist	37%
Meteorologist/reporter	1%
Weather broadcaster	1%
Weekday morning meteorologist	13%
Weekend evening meteorologist	8%
Weekend morning meteorologist	2%
Weekend meteorologist/reporter	2%
Other	3%
Years of experience	
Less than 1 year	1%
1–3 years	11%
4–7 years	17%
8–10 years	7%
More than 10 years	65%
Certifications	
NWA Seal of Approval	29%
AMS Seal of Approval	59%
AMS Certified Broadcast Meteorologist	48%

3. Results

a. Current practices

1) Coverage types and decisions

The first section of this survey more broadly explored how broadcast meteorologists currently communicate NWS tornado and severe thunderstorm products via the technology available to them at their station, and how they use these products to make coverage decisions. For context, the most common coverage decision types include crawling messages, short on-air cut-ins (generally of <5 min), continuous on-air coverage ("wall to wall"), and social media posts (Obermeier et al. 2022). When asked about how NWS products perform within their crawl systems, most respondents indicated that NWS products automatically populate and require little to no effort from the broadcast meteorologist to maintain. This response suggests that most crawl systems have been optimized to handle NWS text products when issued, leaving the broadcaster to focus their efforts on other job duties. Although these systems are designed to run automatically, 64% (n = 101 of 157 total respondents) said they can make manual overrides to their crawls and the information provided by NWS text products if they choose. While most can trust their systems to work properly, 18% of broadcasters (n = 25 of 138 total responses) said they must intervene with their crawl systems with moderate to great frequency to ensure that NWS messages are relayed correctly to their audiences. Nearly all respondents marked that they run crawls for all tornado and severe thunderstorm watches and warnings. A few run a crawl when a significant weather advisory is issued. In the open-ended option, many indicated they also run crawls for flood, winter weather, tropical storm, and nonthunderstorm wind products. One respondent wrote that they would run crawls for "tornado/severe as long as sports coverage isn't on . . . I know, right?," suggesting even the decision to run a crawl comes with some strategizing with regard to programming. Overall, crawl systems provided most respondents an automated and quick way to share NWS watches and warnings with viewers, but some require more intervention than others.

For questions regarding on-air coverage, respondents could select multiple options. Respondents most frequently chose short cut-ins to provide coverage of severe thunderstorm warnings (n = 98) and occasionally to provide coverage of tornado warnings (rather than continuous coverage, n = 78). Several chose to reply with additional explanation in the open-ended option (n = 27). Two common responses emerged regarding severe thunderstorm warnings, including consideration of the intensity of the storm and the population size affected. Broadcasters described using short cut-ins when a storm was "severe enough," "high-end event," "70 mph winds or greater," or "a big enough deal." Like crawling messages, short cuts-ins were also chosen for other hazards such as floods, tropical storms, nonthunderstorm winds and various types of winter weather. One respondent replied, "Don't laugh. Whenever it rains, we have cut-ins at the top of every hour" reflecting that relevant intensity thresholds can apply to a range of storm conditions. Respondents most frequently chose continuous coverage specifically for tornado warnings (n = 127), as compared with severe thunderstorm warnings (n = 14). Twenty-seven respondents provided additional explanations, especially regarding situations in which they might choose continuous coverage for a severe thunderstorm warning (mainly for a derecho). Storm intensity and affected population size were again mentioned frequently in the open-ended option. Also, many explained they would choose continuous coverage for hurricanes and flash floods.

Responses to questions about social media and online practices indicated that Facebook (n = 125), Twitter (n = 114), and station mobile applications (n = 127) were used by the greatest number of respondents to share weather information. Instagram (n = 48) and Snapchat (n = 3) were selected by smaller numbers of respondents. TikTok was not yet listed as an option, as the platform was amidst its rise to widespread popularity in the United States (Laio and Shu 2020). Respondents indicated that they often shared severe thunderstorm and tornado watch and warning information on social media, and only occasionally shared products such as the Storm Prediction Center convective outlooks and mesoscale discussions.

2) Closed captions

During weather coverage, closed captions for severe weather situations are mandated for live broadcasts by the Federal Communications Commission (FCC), including "emergency information accessible to persons who are deaf or hard of hearing, and to persons who are blind or have visual disabilities" (FCC 2021). Captions must be "accurate, synchronous, complete, and properly placed." This mandate means that most coverage types require provision of closed captions during severe weather. Since live severe weather coverage cannot be preplanned in the way scheduled programming or newscasts are, television stations must provide these captions on demand. When asked how closed captioning is handled at their station, respondents indicated several methods. Most said their station provided live closed captioning in English (73%; n = 108), while a limited number said their station provided live Spanish captions (4%; n = 6). In the open-ended option, others indicated that they were not sure how closed captions were handled, or that captions were not available at all. Several respondents indicated that voice to text technology was soon being implemented. For online coverage and social media platforms, nearly 20% (n = 27) indicated that they did not know if closed captions were available at all.

3) Information voids

To better understand how probabilistic information could be used to create a more robust alerting timeline, researchers sought to understand where information voids may exist along the outlook, watch, and warning time scales in the current paradigm. Of 124 responses, most broadcasters felt that "yes" (36%; n = 45) or "maybe" (39%; n = 48) information voids existed. An open-ended follow-up question asked for more explanation, which led to a wide range of perceived issues. Fifteen broadcasters mentioned the lack of updates to a

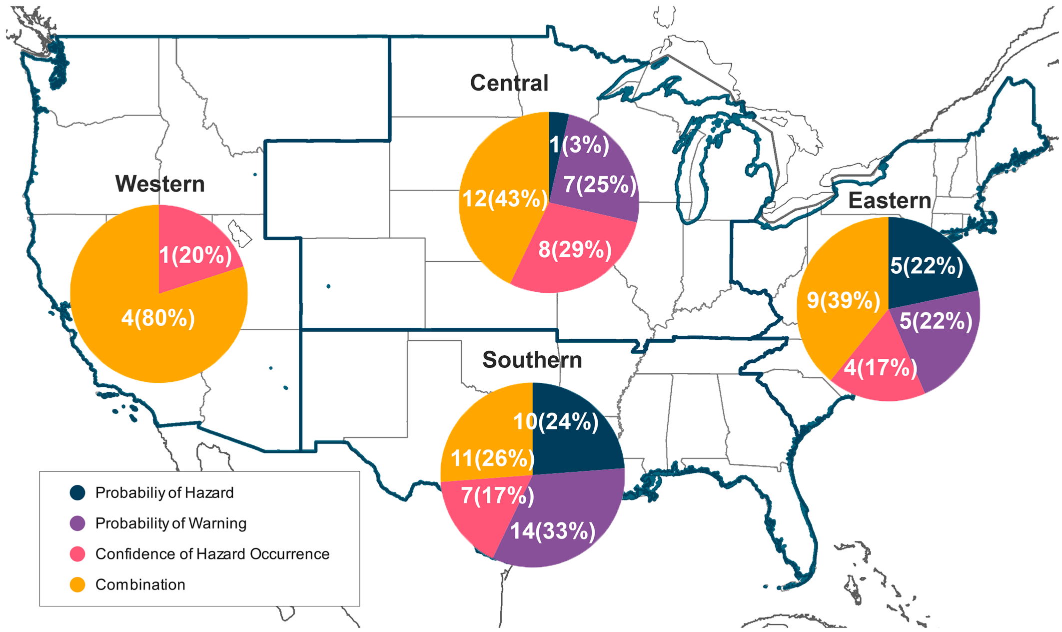

FIG. 2. Preferred definition of PHI (probability of hazard, probability of warning, confidence of hazard occurrence, or a combination of all three options) according to respondents by NWS regions.

warning once issued. A quicker update frequency was desired, with a few stating that waiting 20–30 min for a warning update after tornado or severe thunderstorm warning issuance was too long. In addition, six mentioned the time period between watch issuance and subsequent warnings, in which little to no information was available. One respondent said concerning tornado watches and warnings, "it is either 'potentially within the next 6 hours . . . then it quickly jumps into "it is happening now."' That leaves our viewers who reside in an unsafe structure such as a mobile home almost no time to respond in a beneficial manner." Another respondent said increased communication is lacking in the current system, and helpful information could include "receiving messages such as 'Coordinating with SPC right now . . . watch likely will be issued within the half-hour,' or 'Storm over [county name] starting to show robust mid-level rotation . . . monitoring for potential warning.'" Another respondent mentioned the gap between the convective outlook and the watch/warning timeline, saying "The void is between the SPC outlook (moderate/high) and the watch/warning products." Eight others mentioned confusion and inconsistency, specifically in regard to tags within impact based warnings, saying "tornado possible tag has seemed to cause confusion for viewers despite explanation" and "tags on specific watches/warnings aren't necessarily clear to viewers unless they are watching a newscast and it is explained to them." Regarding inconsistency, respondents were concerned about messaging differences across bordering NWS Weather Forecast Offices (WFOs). For broadcasters whose coverage area spanned multiple WFOs, warning polygon spatial and temporal mismatches were "unique challenges" and difficult "to explain to people why." Other issues included inconsistent warning update frequency, differing color tables in NWS graphics and inconsistency in forecast hazards (e.g., hail or wind was the forecast hazard, but a tornado watch was issued). Four respondents mentioned frequent failures of NWSChat, in which the communication link between the NWS and broadcasters was severed and valuable information was lost.

b. *PHI*

1) DEFINITIONS

For the definition of PHI, respondents could choose from the three options listed in the introduction of this paper (probability of warning, probability of hazard or confidence in hazard occurrence) and could complete an open-ended option to develop a definition in their own words. Respondents were required to choose at least one of the three objective options, although they could choose a combination of the three. To investigate any geographic differences, responses were categorized by corresponding NWS regions for the contiguous United States (Fig. 2). Overall, respondents from all regions preferred a combination of the three definitions. Respondents from the Southern Region chose a single definition more than a combination as compared with other regions; this was especially true for the "probability of warning" and "probability of hazard" options. In addition, 11 respondents used the open-ended option to voice concerns about using probabilities, or to explain why they liked one of the first three options. Some felt that probabilities would be difficult to explain to viewers and would add confusion, writing "throwing in probability . . . is going to muddy the line of communication to the public" and "percentages would create more public confusion." Another

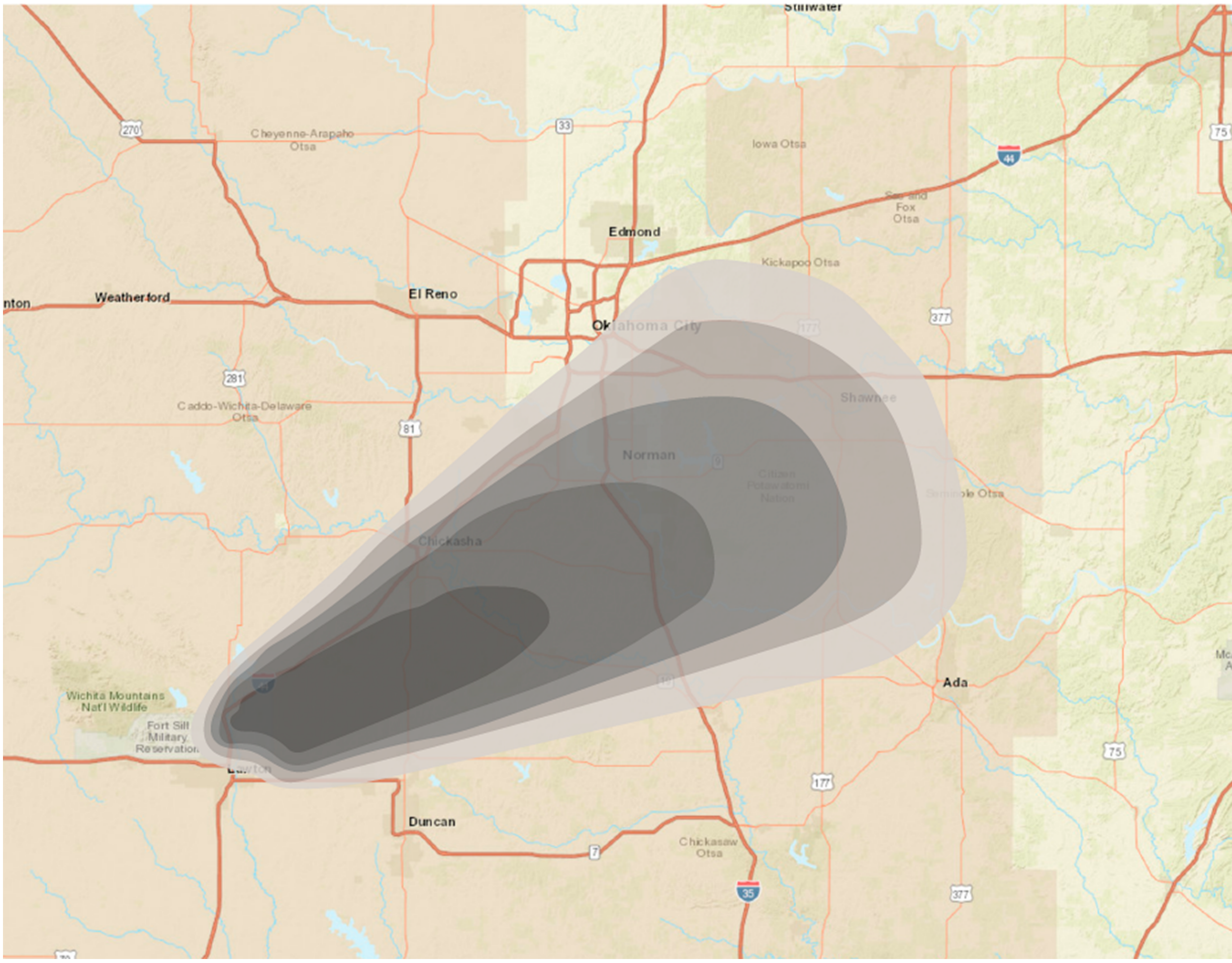

FIG. 3. A static image of PHI displayed in gray tones. Probability increases with darker shades of gray.

respondent suggested qualitative terms, writing that "perhaps a 'low, medium, high' would better serve the public."

2) Thresholds

To explore the idea of thresholds, researchers asked respondents if they might prefer a probabilistic product based upon the Storm Prediction Center's significant severe definitions [wind speeds of 65 kt (33.4 m s^{-1}) or greater; hail 2 in. (5.1 cm) in diameter or larger]. In response, 95 participants preferred a probabilistic product for significant severe hail, and 90 said they would like a probabilistic product for significant severe wind. An open-ended option was also included, and seven respondents mentioned no need for such thresholds because of concerns about confusion and a desire for fewer products. One mentioned that significant severe thresholds are never met in their coverage region, therefore such a product would not be helpful.

3) Nomenclature

Respondents were introduced to an image of PHI in a grayscale monotonic color scheme with no labels (Fig. 3). Researchers asked the respondents what they would call such a product if they were to share it with their viewers. When asked about the nomenclature of PHI, answers contained a diverse range of descriptors (Table 2). Many phrases included the words "severe weather," "probability," "impacts," "threats," "area," "storm," "risk," "tornado(es)" and "cone." Some answers included a form of "storm/severe weather/impact probability," "storm/severe threat" and "storm/severe/impact cone." On the contrary, participants were also asked what words they would not use to describe PHI to their viewers. In contrast to those who answered with the word "probability" in the previous question, 12 respondents of 112 total responses (11%) revealed that using "probability" was undesirable for them. One of these respondents wrote, "Don't use the word probability. The average person with no science background would be very confused." Others wrote, "it would be helpful to give high/medium/low chances to the occurrence, as opposed to labeling each shade with a specific numerical value" and "I think keeping warnings basic and not showing the probabilities during an event will be better for the general public." Eight respondents (7%) indicated they would not share the image or declined to provide a description out of concern for

TABLE 2. The number of times respondents chose to use a word to describe PHI. "Other" indicates words that were used twice or less.

Descriptor	Frequency	Descriptor	Frequency
Probability	29	Zone	7
Impact(s)	18	Track	7
Severe weather	18	Highest	6
Threat	17	Greatest	5
Area	17	Path	4
Storm	16	Warning	4
Risk	15	Hour	4
Tornado(es)	14	Concern	3
Cone	14	Confidence	3
Chance	10	Potential	3
Likely	8	Lighter/lightest	3
Likelihood	7	Hail	3
Hazard	7	Other	30

their viewers' ability to interpret the product. Despite common usage of the word "probability" for many, the word was rejected by others.

4) UPDATE FREQUENCY

For this survey, participants were asked if they might prefer a different update frequency depending on the hazard. Options included 20-, 10-, 5- or 2-min update frequency, for either severe or tornado hazards. Of a total of 98 responses, 60% (n = 58) chose 2-min updates for tornado hazards. Of a total of 97 responses, 45% (n = 43) chose 5-min updates for severe hazards, suggesting that slightly slower update speeds may be more acceptable for wind and hail hazards (Fig. 4). In addition, of 113 total responses, most participants felt comfortable with a mix of human and computer automation for the creation of PHI and subsequent updates (85%; n = 96). Very few preferred the concept of fully algorithm-automated PHI (4%; n = 4) or completely human generated PHI (12%; n = 13).

5) QUANTITATIVE AND QUALITATIVE REPRESENTATION

Survey participants most frequently chose qualitative labeling to share with their viewers (71%; n = 84) as compared with quantitative (13%; n = 15) (Fig. 5). Nineteen participants selected that they would not want to share labels at all. However, 61% (n = 79) indicated they would want to see quantitative data for themselves and their station staff. These answers also echoed some of the thoughts written in response to the questions about definitions and nomenclature, in which broadcasters preferred qualitative labeling out of concern about public reception and explaining explicit probabilities to public viewing audiences.

6) PHI AND WARNING POLYGONS

Participants in this survey were shown three graphical representations, including a warning alone, PHI alone, and PHI and a warning layered with one another (Fig. 6). When asked which they preferred, most chose the image depicting PHI with a warning overlaid (59%; n = 71). Researchers also asked broadcasters if they felt a warning polygon should be drawn spatially within or extending beyond PHI. Four different graphical prototypes of PHI and a warning polygon were shown, including PHI alone (Fig. 7a), a warning drawn completely within PHI (Fig. 7b), a warning drawn with PHI completely contained (Fig. 7c), and a warning drawn with PHI extending beyond (Fig. 7d). Most respondents chose option c [a warning drawn with PHI completely contained (33%; n = 44)] or option b [a warning drawn completely within PHI (34%; n = 46)]. Eighteen respondents (13%) chose a warning drawn with PHI extending beyond the polygon boundaries (Fig. 7d) These responses show that while

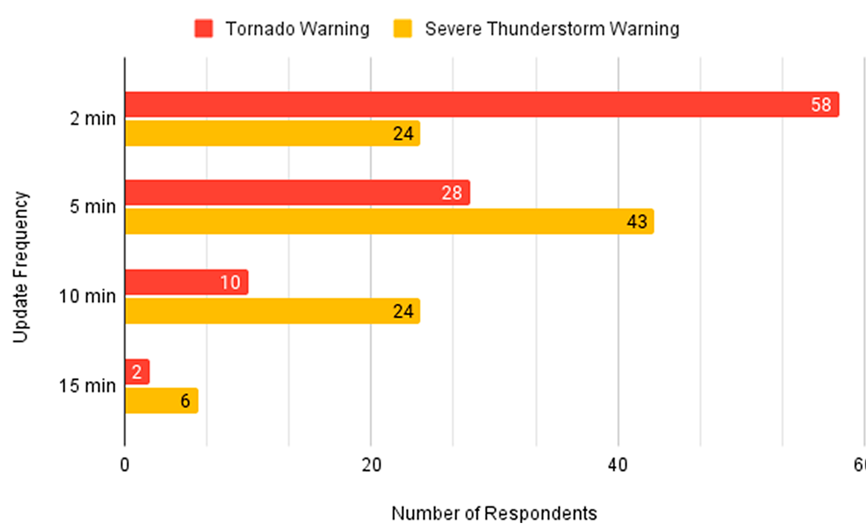

FIG. 4. Preferred update frequency for both tornado and severe thunderstorm warnings according to respondents.

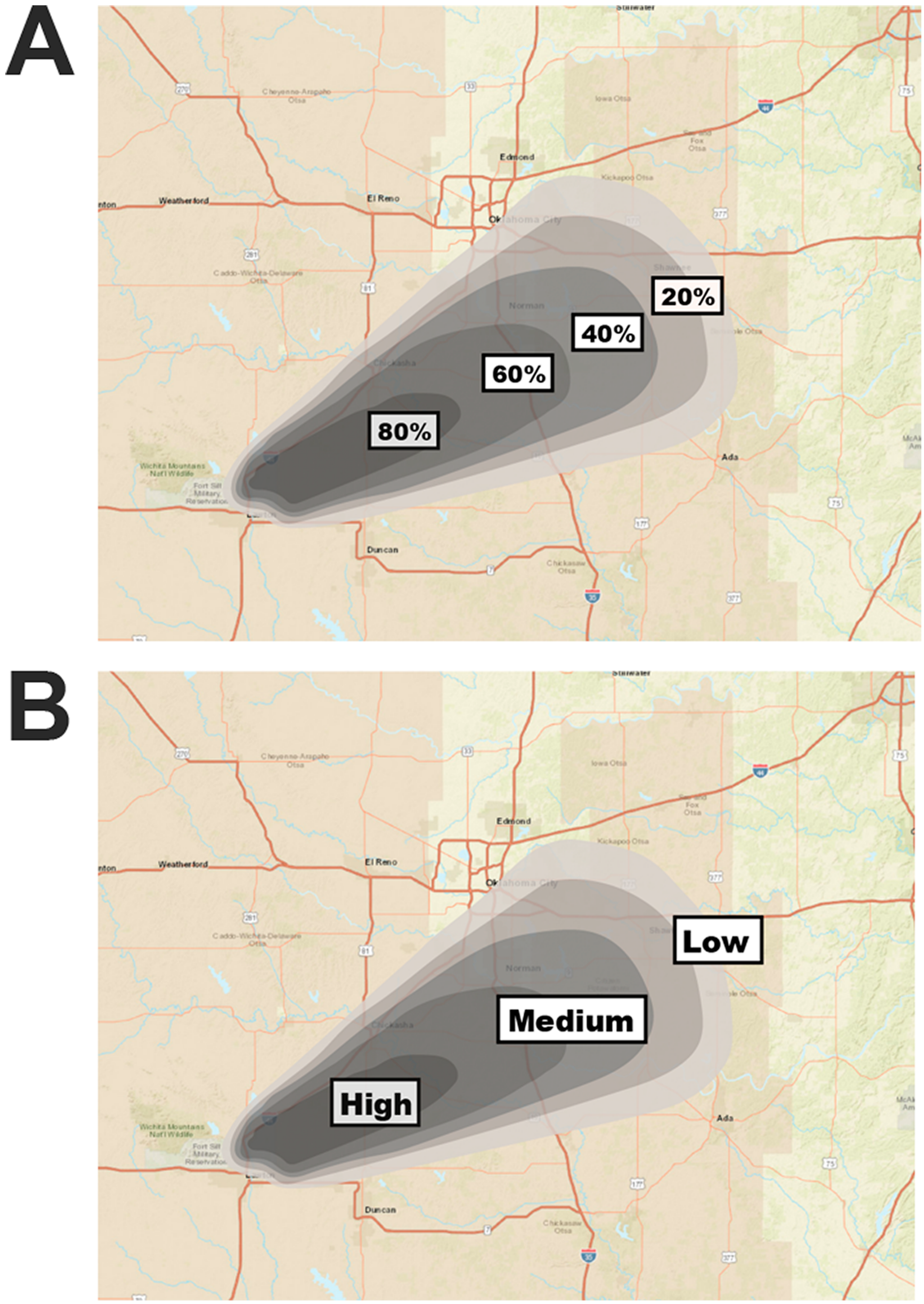

FIG. 5. Static images of PHI displayed with (a) quantitative probability labels and (b) qualitative probability labels. Probability increases with darker shades of gray.

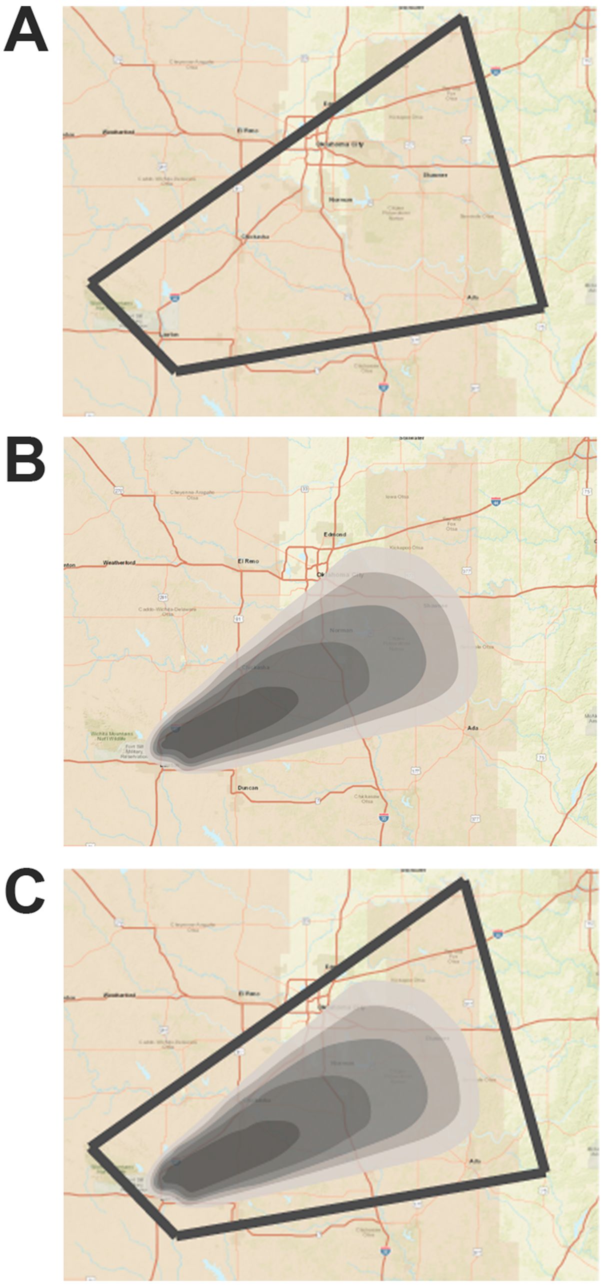

FIG. 6. Static images of PHI and warning polygons. Respondents could choose their preferred display, including (a) a warning polygon alone, (b) PHI alone, or (c) both products displayed with one another.

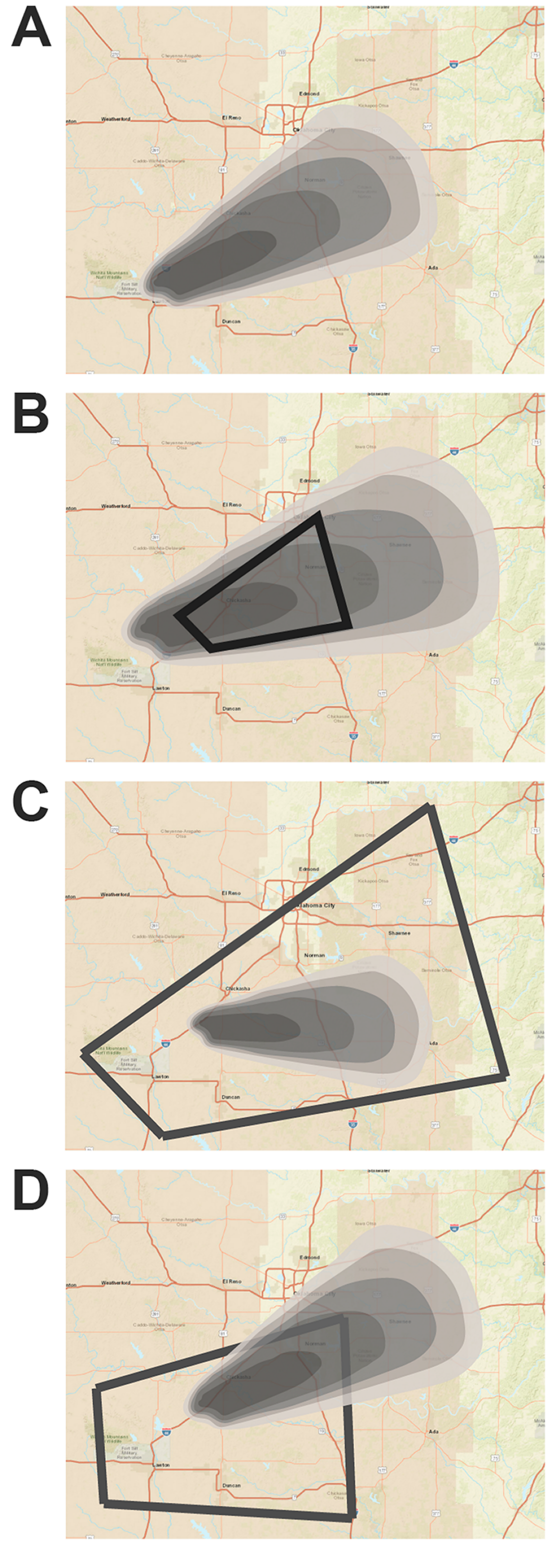

FIG. 7. Static images of PHI and warning polygons. Respondents could choose their preferred display, including (a) PHI alone, (b) a warning polygon within PHI, (c) PHI within a warning polygon, or (d) PHI extending beyond a warning polygon.

most respondents want to see the products simultaneously, they are split on how the representation should look.

4. Discussion and conclusions

This survey outlined in this paper served to provide researchers with a broad pool of responses regarding broadcast meteorologists' current practices regarding their use of NWS outlook, watch, and warning information and to gather more information regarding PHI prototypes to inform the development process. The deployment of this survey represented a component of user research meant to complement prior years of moderated usability testing within the Hazardous Weather Testbed. While many of the findings in this survey are like those found in usability tests, they also provide researchers with the basis to explore additional issues in future research as the FACETs framework continues to move forward.

Several of the respondents to this survey indicated that their crawl systems are not without flaws, requiring attention to remain operational and updated. Current crawl systems may be poorly optimized to handle the rapid updates of PHI, especially for those broadcasters who already experience trouble with their current systems. Similarly, closed captioning requirements could bring about a similar challenge. With increased information from rapidly updating warnings and PHI, the variety of platforms used for severe weather coverage may require more detailed closed captioning or availability of sign language interpretation, including those that are streamed online. Live closed captioning and/or sign language interpreters are the most accurate form; however, these come at an expense not all television stations have the budget to afford. These crawls can be especially important to vulnerable populations in their audience who may rely on these systems as a mode of receiving information (Senkbeil et al. 2021). Although future technologies may help to mitigate the problems described here, developers should strive to keep current tools and compatibility in mind.

Many respondents indicated that NWS warnings are a cue for them to provide on-air coverage, in the form of short cut-ins or continuous coverage. However, open-ended answers from respondents indicated that their coverage decisions are more complex. Factors such as hazard intensity and the geographic characteristics of their market area influence their decisions, and a "one size fits all" recipe for coverage decisions does not exist. Regarding watches and warnings, too little information was available specifically between the time of watch issuance and subsequent warnings. Once warnings have been issued, many respondents mentioned a lack of updates or inconsistent updates to the warnings. Another concern was inconsistency in messaging from neighboring NWS WFOs. Researchers have recognized this issue and continue to take steps to help mitigate such challenges with future updates to the warning paradigm (Trujillo-Falcón et al. 2022b).

Respondents were split on their preferred definition of PHI and had concerns about how their viewership may understand or interpret the meaning of PHI. While many respondents seemed comfortable with the idea of PHI, some were wary about how their audiences would understand PHI. Their concerns were evident in their responses regarding the nomenclature and labeling of PHI, as some were unsure as to whether the word probability should be used at all. However, the overall body of literature indicates that the inclusion of probabilistic information in weather forecasts does improve audience decision-making, and according to the "living systematic review" conducted by Ripberger et al. (2022), "Nearly all of the studies we review indicate that people make better decisions, have more trust in information, and/or display more understanding of forecast information when forecasters use probability information in place of deterministic statements." It may be of interest for researchers to learn more about why some broadcasters feel opposed to the explicit use of probability for severe weather hazards. Regarding labeling, literature indicates that messaging that includes a mixture of quantitative and qualitative probabilities is recommended (Budescu et al. 2014; Kox et al. 2015; Lenhardt et al. 2020; Ripberger et al. 2022). The qualitative example of PHI shown in the survey outlined in this paper included a simple three point Likert scale, similar to those tested in the Hazardous Weather Testbed. This prototype was meant to elicit a simplistic preference when compared with a quantitative example, however, additional qualitative probability designs need to be developed and tested more thoroughly to gain a better understanding of end-user communication needs.

Usability studies have repeatedly shown that the relationship between PHI and the warning polygons issued by NWS forecasters matters, and developers should carefully consider the interplay between the two products. While most respondents to this survey indicated they would show both PHI and a warning polygon simultaneously, they were split on how the representation between the two should appear. As described in the introduction section, the boundary and shape of a warning polygon can have an impact on perception. Less is known about how the layering of a probabilistic plume could influence or change these perceptions. Early research suggests that different abstractions of the two products does have an impact on sheltering decisions with members of the public (Qin et al. 2023). More systematic research on this topic is essential, in regard to both core partners and the public.

A central finding from this survey is that refined, yet user-driven, options should be a key component of PHI. Because of factors such as the diversity of coverage platforms, technology, geographical differences, and local viewership populations and culture, PHI should not be a one-size-fits-all product. While best practices and a refined set of options are foundational, broadcasters should have the ability to choose what to share with their viewers concerning probabilistic hazard criteria thresholds, nomenclature, qualitative/quantitative probabilities, and update frequencies. For example, developers could give users the ability to share both qualitative and quantitative labels rather than being limited to one or the other (Krocak et al. 2022). Also, the needs of other core partners, which may be different from broadcasters' needs, must also be considered in development. This consideration further compounds the need for a customizable PHI.

Limitations

Although this survey provided a larger dataset of feedback from broadcast meteorologists as compared with aforementioned usability studies in the Hazardous Weather Testbed, it is not without limitations. From a research perspective, the broadcast meteorologist community can be somewhat difficult to access due to the nature of private industry. For this reason, it is challenging to ascertain the demographics of the overall population and therefore create a baseline for the demographics of the participants in this survey. This survey appeared to have fewer respondents from western states. Designated market areas tend to be much larger with western extent, meaning fewer broadcast meteorologists cover these regions, but it would seem the number of respondents is indeed lower than the overall population. Despite less severe convective weather in this region, understanding coverage practices and user needs is still important for PHI development for hazards such as floods and wildfires.

In addition, this survey captured few Spanish-speaking broadcast meteorologists, who can provide more insights regarding current practices and PHI. Trujillo-Falcón et al. (2022a) noted that significant language inequities exist among U.S. Spanish speakers when receiving warning information due to inconsistent communication among the bilingual weather enterprise. Future survey work and studies in broadcast meteorologist practices will need to focus on these groups more specifically to address language and cultural inequities. Another limitation includes the static nature of the images in the questions regarding PHI and its representation with the warning polygon. Since PHI is designed to rapidly update, a single image does not capture how the plume will move in relation to the warning over time.

As FACETs paradigm research continues, investigators must emphasize the importance of the end user and the iterative nature of the development process. Using complementary methods, such as surveys and usability studies, allows researchers to refine new products and increase the likelihood of operational success for all users. Through such research with NWS core partners and the public, new products can be designed to better serve users and optimize outcomes.

Acknowledgments. Funding was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA21OAR4320204, U.S. Department of Commerce. The authors also thank Makenzie Krocak, Caroline Carithers, and Nate Johnson for pretesting this survey, as well as three internal reviewers who provided helpful comments.

Data availability statement. Because of privacy and ethical concerns outlined in human-subjects research guidelines and adherence to confidentiality agreements with the participants of this study, neither the data nor additional information about the data or the study participants can be made available.

REFERENCES

AWARN, 2022: Advanced warning and response network. Accessed 27 April 2022, https://awarn.org.

Bitterman, A., M. Krocak, J. Ripberger, C. Silva, H. Jenkins-Smith, S. Ernst, and S. Stormer, 2022: Maintaining the series: Public reception, understanding, and responses to severe weather forecasts and warnings in the contiguous United States. Harvard Dataverse, accessed 1 August 2022, https://doi.org/10.7910/DVN/TD5DGD.

Braun, V., and V. Clarke, 2006: Using thematic analysis in psychology. *Qual. Res. Psychol.*, **3**, 77–101, https://doi.org/10.1191/1478088706qp063oa.

Budescu, D. V., H.-H. Por, S. B. Broomell, and M. Smithson, 2014: The interpretation of IPCC probabilistic statements around the world. *Nat. Climate Change*, **4**, 508–512, https://doi.org/10.1038/nclimate2194.

Calhoun, K. M., and Coauthors, 2018: Cloud-to-ground lightning probabilities and warnings within an integrated warning team. *Special Symp. on Impact-Based Decision Support Services*, Austin, TX, Amer. Meteor. Soc., 4.4, https://ams.confex.com/ams/98Annual/webprogram/Paper329888.html.

——, K. L. Berry, D. M. Kingfield, T. Meyer, M. J. Krocak, T. M. Smith, G. Stumpf, and A. Gerard, 2021: The experimental warning program of NOAA's Hazardous Weather Testbed. *Bull. Amer. Meteor. Soc.*, **102**, E2229–E2246, https://doi.org/10.1175/BAMS-D-21-0017.1.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, and D. T. Lindsey, 2014: An empirical model for assessing the severe weather potential of developing convection. *Wea. Forecasting*, **29**, 639–653, https://doi.org/10.1175/WAF-D-13-00113.1.

Cranford, A., 2018: Women weathercasters: Their positions, education, and presence in local TV. *Bull. Amer. Meteor. Soc.*, **99**, 281–288, https://doi.org/10.1175/BAMS-D-16-0317.1.

Ernst, S. R., 2020: Colorful language: Investigating the interpretation of the Storm Prediction Center's convective outlook by broadcast meteorologists and the U.S. public. M.S. thesis, Dept. of Meteorology, University of Oklahoma, 124 pp.

FCC, 2021: Access to emergency information on television. FCC, accessed 14 October 2022, https://www.fcc.gov/consumers/guides/accessibility-emergency-information-television.

Harrison, D. R., and C. D. Karstens, 2017: A climatology of operational storm-based warnings: A geospatial analysis. *Wea. Forecasting*, **32**, 47–60, https://doi.org/10.1175/WAF-D-15-0146.1.

Hatfield, K., S. F. Piltz, E. J. Calianese Jr., J. M. Frederick, J. B. McGavock, and M. E. Plate, 2018: Implementation of the WFO Tulsa vision for local impact-based decision support services. *Special Symp. on Impact-Based Decision Support Services*, Austin, TX, Amer. Meteor. Soc., 1.3, https://ams.confex.com/ams/98Annual/webprogram/Manuscript/Paper329294/1_3_Hatfield_ExtendedAbstract.pdf.

——, and Coauthors, 2023: Turning severe weather warning probabilities into action: NWS Tulsa efforts to better bridge the watch-to-warning gap. *Special Symp. on Forecasting a Continuum of Environmental Threats (FACETs)*, Denver, CO, Amer. Meteor. Soc., J3.4, https://ams.confex.com/ams/103ANNUAL/meetingapp.cgi/Paper/414104.

Karstens, C., T. M. Smith, K. M. Calhoun, A. J. Clark, C. Ling, G. J. Stumpf, and L. P. Rothfusz, 2014: Prototype tool development for creating probabilistic hazard information for severe convective weather. *Second Symp. on Building a Weather-Ready Nation: Enhancing Our Nation's Readiness, Responsiveness, and Resilience to High Impact Weather Events*, Atlanta, GA, Amer. Meteor. Soc., 2.2, https://ams.confex.com/ams/94Annual/webprogram/Manuscript/Paper241549/Karstens_AMS_2014.pdf.

——, and Coauthors, 2015: Evaluation of a probabilistic forecasting methodology for severe convective weather in the 2014 Hazardous Weather Testbed. *Wea. Forecasting*, **30**, 1551–1570, https://doi.org/10.1175/WAF-D-14-00163.1.

——, and Coauthors, 2018: Development of a human–machine mix for forecasting severe convective events. *Wea. Forecasting*, **33**, 715–737, https://doi.org/10.1175/WAF-D-17-0188.1.

Klockow-McClain, K. E., R. A. McPherson, and R. P. Thomas, 2020: Cartographic design for improved decision-making: Trade-offs in uncertainty visualization for tornado threats. *Ann. Amer. Assoc. Geogr.*, **110**, 314–333, https://doi.org/10.1080/24694452.2019.1602467.

Kolakowski, C., K. Berry, A. Gerard, H. Obermeier, J. T. Ripberger, P. A. Campbell, T. C. Meyer, and K. E. Klockow-McClain, 2019: Broadcast meteorologist use of social media in the 2018 Hazardous Weather Testbed probabilistic hazard information project. *Ninth Conf. on Transition of Research to Operations*, Phoenix, AZ, Amer. Meteor. Soc., 8B.5, https://ams.confex.com/ams/2019Annual/recordingredirect.cgi/oid/Recording50929/paper351124_1.mp4.

Kox, T., L. Gerhold, and U. Ulbrich, 2015: Perception and use of uncertainty in severe weather warnings by emergency services in Germany. *Atmos. Res.*, **158–159**, 292–301, https://doi.org/10.1016/j.atmosres.2014.02.024.

Krocak, M. J., J. T. Ripberger, C. L. Silva, H. C. Jenkins-Smith, S. Ernst, A. Bell, and J. Allan, 2020: Measuring change: Public reception, understanding, and responses to severe weather forecasts and warnings in the contiguous United States, version 2. University of Oklahoma Center for Risk and Crisis Management, Harvard Dataverse, accessed 1 August 2022, https://doi.org/10.7910/DVN/EWOCUA.

——, ——, C. Silva, H. Jenkins-Smith, S. Ernst, A. Gaviria-Pabón, A. Forney, and A. Bitterman, 2021: Continuing the baseline: Public reception, understanding, and responses to tropical cyclone forecasts and warnings in the contiguous United States. University of Oklahoma Center for Risk and Crisis Management, Harvard Dataverse, accessed 1 August 2022, https://doi.org/10.7910/DVN/QYZLSO.

——, ——, S. Ernst, C. L. Silva, and H. C. Jenkins-Smith, 2022: Exploring the differences in SPC convective outlook interpretation using categorical and numeric information. *Wea. Forecasting*, **37**, 303–311, https://doi.org/10.1175/WAF-D-21-0123.1.

Laio, R., and C. Shu, 2020: TikTok's epic rise and stumble. TechCrunch, accessed 19 May 2023, https://techcrunch.com/2020/11/26/tiktok-timeline/.

Lenhardt, E. D., R. N. Cross, M. J. Krocak, J. T. Ripberger, S. R. Ernst, C. L. Silva, and H. C. Jenkins-Smith, 2020: How likely is that chance of thunderstorms? A study of how National Weather Service forecast offices use words of estimative probability and what they mean to the public. *J. Oper. Meteor.*, **8**, 64–78, https://doi.org/10.15191/nwajom.2020.0805.

Lindell, M. K., 2020: Improving hazard map comprehension for protective action decision making. *Front. Comput. Sci.*, **2**, 27, https://doi.org/10.3389/fcomp.2020.00027.

Meyer, T. C., K. Berry, K. M. Calhoun, H. Obermeier, A. Campbell, J. Wolfe, and K. E. Klockow-McClain, 2019: Visualization and communication of probabilistic lightning hazard information for a broad spectrum of end users. *Ninth Conf. on the Meteorological Application of Lightning Data*, Phoenix, AZ, Amer. Meteor. Soc., 6.3, https://ams.confex.com/ams/2019Annual/recordingredirect.cgi/oid/Recording50549/paper349844_1.mp4.

Nemunaitis-Berry, K. L., and H. M. Obermeier, 2017: Broadcast meteorologist decision-making in a probabilistic hazard information project. *Bull. Amer. Meteor. Soc.*, **98**, 2045–2047, https://doi.org/10.1175/BAMS_9810_2039-2052_Nowcast.

——, H. Obermeier, S. A. Jasko, D. LaDue, C. D. Karstens, G. M. Eosco, A. Gerard, and L. Rothfusz, 2017: Broadcast meteorologist decision making in the 2016 Hazardous Weather Testbed probabilistic hazard information project. *Fourth Conf. on Weather Warnings and Communication*, Kansas City, MO, Amer. Meteor. Soc., 1.4, https://ams.confex.com/ams/45BC4WXCOMM/webprogram/Paper318190.html.

——, ——, K. M. Calhoun, T. C. Meyer, K. E. Klockow-McClain, D. LaDue, Z. Stanford, and A. Gerard, 2020: Incorporating end users in Hazardous Weather Testbed experiments. *10th Conf. on Transition of Research to Operations*, Boston, MA, Amer. Meteor. Soc., 2.3, https://ams.confex.com/ams/2020Annual/webprogram/Paper363946.html.

Obermeier, H., K. L. Nemunaitis-Berry, S. A. Jasko, D. LaDue, C. Karstens, G. M. Eosco, A. Gerard, and L. Rothfusz, 2017: Broadcast meteorologist decision making in the 2016 hazardous weather test probabilistic hazard information project. *12th Symp. on Societal Applications: Policy, Research and Practice*, Seattle, WA, Amer. Meteor. Soc., 4.3, https://ams.confex.com/ams/97Annual/webprogram/Paper314239.html.

——, K. Nemunaitis-Berry, K. E. Klockow, C. D. Karstens, A. Gerard, and L. P. Rothfusz, 2018: Broadcast meteorologist decision-making in the 2017 Hazardous Weather Testbed probabilistic hazard information project. *13th Symp. on Societal Applications: Policy, Research and Practice*, Austin, TX, Amer. Meteor. Soc., 8.4, https://ams.confex.com/ams/98Annual/webprogram/Paper329820.html.

——, K. L. Nemunaitis-Berry, K. E. Klockow-McClain, T. C. Meyer, P. A. Campbell, A. Gerard, and C. Kolakoski, 2019: Communicating probabilistic hazard information: Broadcast meteorologists in the 2018 Hazardous Weather Testbed. *47th Conf. on Broadcast Meteorology/Fifth Conf. on Weather Warnings and Communication*, San Diego, CA, Amer. Meteor. Soc., J7.6, https://ams.confex.com/ams/47BC5WxComm/recordingredirect.cgi/oid/Recording54984/paper358620_1.mp4.

——, ——, ——, ——, ——, ——, J. E. Trujillo, and C. Carithers, 2020: Communicating probabilistic hazard information: Broadcast meteorologists in the 2018-19 Hazardous Weather Testbed. *15th Symp. on Societal Applications: Policy, Research and Practice*, Boston, MA, Amer. Meteor. Soc., 12A.4, https://ams.confex.com/ams/2020Annual/meetingapp.cgi/Paper/368267.

——, 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.

Perkins, D. R., K. Timm, T. Myers, and E. Maibach, 2020: Broadcast meteorologists' views on climate change: A state-of-the-community review. *Wea. Climate Soc.*, **12**, 249–262, https://doi.org/10.1175/WCAS-D-19-0003.1.

Qin, C., S. Joslyn, S. Savelli, J. Demuth, R. Morss, and K. Ash, 2023: The impact of probabilistic tornado warnings on risk perceptions and responses. *J. Exp. Psychol. Appl.*, https://doi.org/10.1037/xap0000486, in press.

Reed, J. R., and J. C. Senkbeil, 2020: Perception and comprehension of the extended forecast graphic: A survey of broadcast meteorologists and the public. *Bull. Amer. Meteor. Soc.*, **101**, E221–E236, https://doi.org/10.1175/BAMS-D-19-0078.1.

Ripberger, J., A. Bell, A. Fox, A. Forney, W. Livingston, C. Gaddie, C. Silva, and H. Jenkins-Smith, 2022: Communicating probability information in weather forecasts: Findings and recommendations from a living systematic review of the research literature. *Wea. Climate Soc.*, **14**, 481–498, https://doi.org/10.1175/WCAS-D-21-0034.1.

Robertson, M. M., and K. K. Droegemeier, 1990: NEXRAD and the broadcast weather industry: Preparing to share the technology. *Bull. Amer. Meteor. Soc.*, **71**, 14–18, https://doi.org/10.1175/1520-0477(1990)071<0014:NATBWI>2.0.CO;2.

Rothfusz, L. P., R. Schneider, D. Novak, K. Klockow-McClain, A. E. Gerard, C. Karstens, G. J. Stumpf, and T. M. Smith, 2018: FACETs: A proposed next-generation paradigm for high-impact weather forecasting. *Bull. Amer. Meteor. Soc.*, **99**, 2025–2043, https://doi.org/10.1175/BAMS-D-16-0100.1.

Senkbeil, J. C., D. J. Griffin, K. Sherman-Morris, J. Saari, and K. Brothers, 2021: Improving tornado warning communication for deaf and hard of hearing audiences. *J. Oper. Meteor.*, **9**, 18–35, https://doi.org/10.15191/nwajom.2021.0902.

Silva, C. L., J. T. Ripberger, H. C. Jenkins-Smith, and M. Krocak, 2017: Establishing a baseline: Public reception, understanding, and responses to severe weather forecasts and warnings in the contiguous United States. University of Oklahoma, Center for Risk and Crisis Management Reference Rep., 31 pp., https://www.e-education.psu.edu/meteo3/sites/www.e-education.psu.edu.meteo3/files/images/lesson13/WX17-Reference-Report.pdf.

——, ——, ——, ——, and W. W. Wehde, 2018: Refining the baseline: Public reception, understanding, and responses to severe weather forecasts and warnings in the contiguous United States. Center for Risk and Crisis Management Reference Rep., 29 pp.

——, ——, ——, ——, S. Ernst, and A. Bell, 2019: Continuing the baseline: Public reception, understanding, and responses to severe weather forecasts and warnings in the contiguous United States. Center for Risk and Crisis Management Reference Rep., 33 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.

——, and Coauthors, 2022a: ¿Aviso o alerta? Developing effective, inclusive, and consistent watch and warning translations for U.S. Spanish speakers. *Bull. Amer. Meteor. Soc.*, **103**, E2791–E2803, https://doi.org/10.1175/BAMS-D-22-0050.1.

——, J. Reedy, K. E. Klockow-McClain, K. L. Berry, G. J. Stumpf, A. V. Bates, and J. G. LaDue, 2022b: 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.

Wilson, K. A. P., and Coauthors, 2023: Five years of user-focused warn-on-forecast research. *Special Symp. on Forecasting a Continuum of Environmental Threats (FACETs)*, Denver, CO, Amer. Meteor. Soc., 5.1, https://ams.confex.com/ams/103ANNUAL/meetingapp.cgi/Paper/417823.