BAMS Article

Conceptualizing Confidence

A Multisited Qualitative Analysis in a Severe Weather Context

Jen Henderson, Jennifer Spinney, and Julie L. Demuth

ABSTRACT: Confidence is a concept important to weather prediction, shaping how risk information is created, shared, understood, and acted upon. For forecasters in the National Weather Service (NWS) and their partners in public safety, confidence is central to their work, appearing frequently during their decision support services. While confidence has been examined in a variety of literatures, it is often addressed simplistically or as one of many variables in a study. It is rarely the object of study in and of itself, even less so in a naturalistic setting like an operational environment. To build a more robust knowledge of confidence and its many dimensions, we conducted a multisited ethnography of three interrelated sites central to tornado prediction and information dissemination, leading up to and during a cool-season tornado event. In partnership with collaborators from the NWS and emergency management, we simultaneously deployed to a National Center, a local Weather Forecast Office, and an emergency management office. This article explicates confidence from multiple social science theories, considering the scientific, data-based roots of confidence, as well as its affective, relational, and procedural origins. Our results show that confidence emerges in varied and complex ways and at different scales. Confidence can indicate one's assessment of evidence and agreement (or lack) of it, beliefs about partners' future behavior based on past experiences, and ritual interactions between offices that create patterned expectations. We argue for a more robust interdisciplinary analysis of confidence given how it shapes weather-related policies and practices, technologies, and communication strategies.

KEYWORDS: Social Science; Forecasting; Decision making; Emergency response; Risk assessment; Societal impacts

https://doi.org/10.1175/BAMS-D-22-0137.1

Corresponding author: Jen Henderson, jen.henderson@ttu.edu

In final form 18 November 2022

AFFILIATIONS: Henderson—Texas Tech University, Lubbock, Texas; **Spinney**—York University, Toronto, Ontario, Canada; **Demuth**—National Center for Atmospheric Research, Boulder, Colorado

Confidence is ubiquitous in human experience. It helps us judge how much credence to place in something or how we feel about our own abilities. The notion also is prevalent in the weather forecasting domain, often framed as a need to explicitly characterize forecast confidence and communicate it to end users. The concept of confidence is of longstanding interest, as indicated by Williams (1951), who conducted an experiment requiring forecasters to assign a number representing a probability statement to each of their precipitation forecasts, as a way to elicit "the degree of confidence [a forecaster] has that the forecast will verify." Efforts like this, in part, reflect the belief that forecasts must be useful to be good (Murphy 1993), and that confidence plays a key role in this utility. Fischhoff and MacGregor (1982) explicitly stated this, indicating that "forecasts have little value to decision makers unless it is known how much confidence to place in them. Those expressions of confidence have, in turn, little value unless forecasters are able to assess the limits of their own knowledge accurately." These ideas remain relevant today, as evidenced by the National Weather Service's (NWS's) 2019–22 Strategic Plan, which includes an objective to "improve the quantification of confidence, specificity, and potential impacts within forecasts and warnings" as a means of providing "better information for better decisions."

NWS forecasters have long expressed their confidence qualitatively in freeform text forecast discussions, written in local Weather Forecast Offices (WFO) and National Centers. Additionally, NWS forecasters, broadcast meteorologists, and others have experimented with conveying forecast confidence ordinally, either qualitatively (e.g., low, medium, high) or using a numeric scale (Demuth et al. 2009). Correspondingly, recent research by Padilla et al. (2021) experimentally investigated the effects of qualitatively conveying forecaster confidence as low, medium, or high on users' decisions. Examples of confidence visualizations taken from NWS WFO websites are shown in Figs. 1 and 2.

Moreover, several studies with members of the public have examined their confidence in forecast information (e.g., Kox and Thieken 2017; Demuth et al. 2011; Lazo et al. 2009; Joslyn and Savelli 2010; Hogan Carr et al. 2016; Su et al. 2021). Confidence in this context is typically measured through single, Likert-scale items and as one of several dependent variables of interest rather than a focal variable of interest. Although this research contributes to how confidence is conceptualized, it also reveals an assumption that confidence solely pertains to scientific data, that it ought to be characterized systematically and, ideally, quantitatively, and that it is transferable in a simplified way to decision-makers.

In this article, we suggest that there needs to be broader consideration of confidence, including from where and how it emerges and is shaped, how it manifests and operates, and how and for whom it has an effect. In doing so, we consider the scientific, data-based roots of confidence, but we also consider the affective, relational, and procedural origins. In short, we explore confidence as having a complex "social life" (Duguid 2005). We contend that such work is essential given the centrality of confidence to the prediction and communication of, as well as the responses to, weather risks.

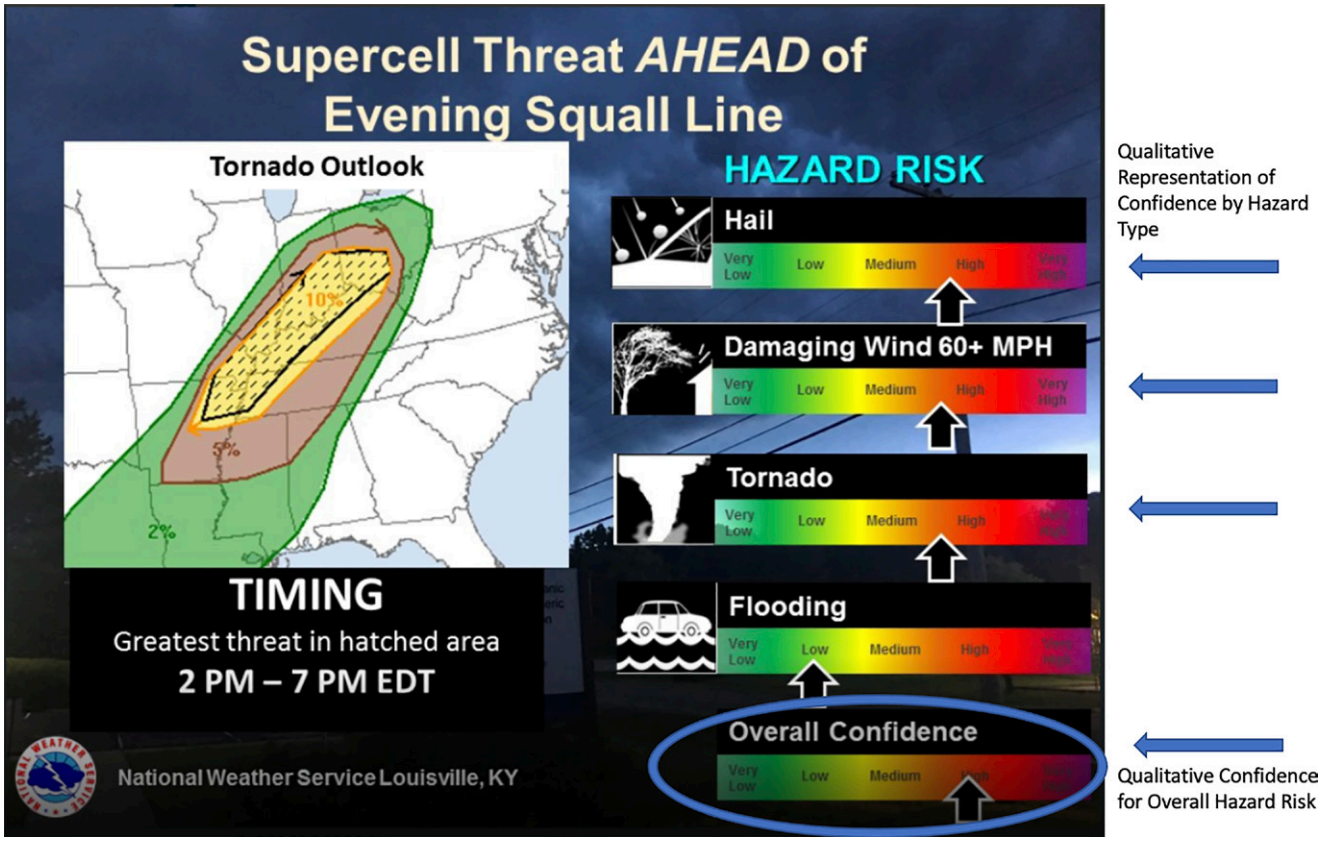
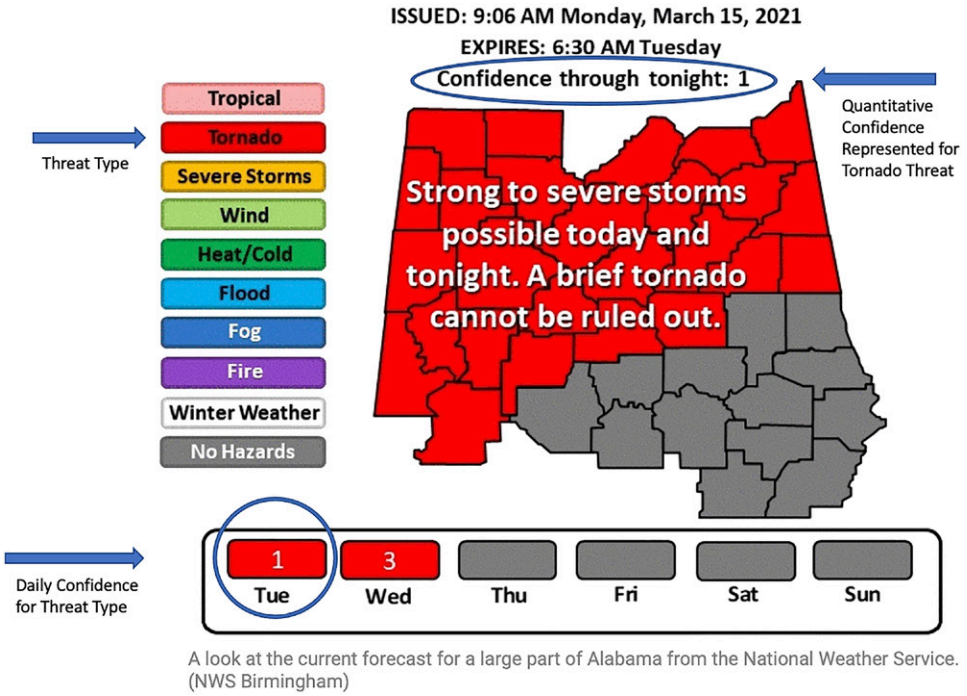


Fig. 1. In this NWS graphic, confidence is depicted qualitatively as different colors corresponding to increasing risk levels, from green to magenta, and the forecasters' confidence in each type of hazard risk, as indicated by black arrows. Annotations with blue arrows, blue ovals, and text were made by the authors to highlight where and how confidence is represented for this graphic.



A look at the current forecast for a large part of Alabama from the National Weather Service. (NWS Birmingham)

Fig. 2. In this NWS graphic, confidence is depicted quantitatively with a number assigned to the dominant threat type, in this case a tornado threat represented by the color red. For each day of the week, the threat is assigned a number representing their confidence (i.e., Tuesday is a 1; Wednesday is a 3), although it is unclear what the scale range is. Annotations with blue arrows, blue ovals, and text were made by the authors to highlight where and how confidence is represented for this graphic.

We situate our study of confidence amid diverse social science theories and frameworks to illustrate multiple expressions and functions of the concept. We do so by presenting brief, empirical examples of confidence based on multisited ethnographic field work of a cool-season tornado event in the southeastern United States with three groups of experts who are central to the functions of weather prediction and response: NWS forecasters at both a National Center and a local WFO, and emergency management partners (EMs) in the county warning area (CWA) of the WFO studied. To our knowledge, research has not been conducted that examines how confidence plays out in real time for severe weather among and between groups of experts responsible for public safety.

Drawing on literatures from science and technology studies, sociology, and risk assessment and risk communication, we report on examples of confidence that illustrate the diversity of ways confidence might be characterized. Further, we build knowledge about how it is mobilized in different decision-making spaces. We argue that we need increased research and analytic attention to the concept of confidence, given the many ways it is unaccounted for in weather warning systems and community disaster resilience. Such explication will help provide insight into how confidence might be leveraged to facilitate clearer communication of risk and to better prepare people for and potentially mitigate against harm.

In what follows, we provide background context for the study, details about our data collection and methods, and theories that help explicate examples of how confidence emerged for the cool-season tornado event we studied.

Background

Impact-based decision support services. Impact-based decision support services (IDSS) is an NWS initiative that has as its goal information support for partners in public safety through a host of activities to ensure effective communication, from webinars and briefing packets to tabletop exercises and workshops (National Weather Service 2019). IDSS is practiced through daily activities of engagement with partners over time but is particularly important leading up to and during severe weather (Cavanaugh et al. 2016). A central element to relationships developed between NWS and partners is their practices for sharing of evolving weather risk information. In part, these are built on knowledge of one another's roles during severe weather, including decisions made at varying decision thresholds and information needs. Importantly, communication in support of IDSS involves and often hinges on discussions about forecaster confidence about which hazards will occur, when, and at what magnitude or severity. While IDSS is often focused on local NWS WFO forecasters and their local partners, we extend this framework to include NWS forecasters in National Centers, like the Weather Prediction Center and the Storm Prediction Center, given the importance of their work to the provision of risk information. This expansive view of IDSS allows concepts like confidence to be studied more comprehensively across communication timelines and decision spaces and helps researchers better characterize interactions between expert groups and the many propagations of confidence that may arise.

Analytic partners. We selected three sites for our ethnographic field work: an NWS National Center, an NWS WFO, and an Office of Emergency Management (OEM) in the CWA of the WFO being studied. These groups comprise part of the integrated forecast and warning system (Doswell et al. 1999), sometimes referred to as an integrated warning team (e.g., Morris et al. 2008), and they frequently engage in practices of IDSS. The interconnections among these three groups and their critical roles for informing and protecting society about severe weather risks guided our selection for studying how confidence emerges within and among them. Due to IRB privacy requirements, names, places, and other identifying information have been anonymized for this manuscript.

For social scientists, these expert groups are likewise critical to knowledge construction. To that end, we consider our collaborators our "analytic partners," given that they codeveloped the research approach with us by guiding research questions, the types of phenomena studied, and timing for observations. This framing mirrors the epistemic orientation of community-based participatory action research (CBPR), which makes communities equal partners in problem definition and recognizes the value of multiple knowledges throughout the research process (Belone et al. 2016; Henderson and Liboiron 2019). CBPR scholarship seek to transform the scientific enterprise through improvements in the "*rigor*, *relevance* and *reach* of science" (Balazs and Morello-Frosch 2013). Important to this point, our analytic partners provided input throughout the research and analysis process, including input on drafts of this article, thus shaping how research might serve the operational needs for those involved.

Partner roles. The mission of the NWS National Center of study is to provide nationwide timely and accurate forecast products regarding the threat of severe convective weather hazards. These forecasters issue a suite of products, including convective outlooks, which spatially depict areas where severe weather could occur and which are issued beginning multiple days before and up to the day that severe weather threatens. Within 3 days of the threat of severe weather, the convective outlooks issued are categorical. The five categories—Marginal, Slight, Enhanced, Moderate, and High—are associated with a probability of the occurrence of tornadoes, severe wind (58 mph or greater; 1 mph $\approx$ 0.45 m s^{-1}), and severe hail (1 in. or greater; 1 in. = 2.54 cm). The day of a severe weather threat, the National Center also may issue mesoscale convective discussions, severe thunderstorm watches, and/or tornado watches, as warranted. These forecasters regularly communicate directly with local WFO forecasters when severe weather threatens.

NWS forecasters who work in one of 122 WFOs across the United States provide local forecasts and warnings for weather for their respective CWAs, including the daily "blue sky" forecasts and potential hazardous weather. When severe weather is a possibility, WFO forecasters assess these threats in the days leading up to a hazard and, in collaboration with National Centers and local partners, issue risk information and products with specific information about the timing, magnitude, severity, and location of these threats. For severe convective weather hazards, WFO forecasters use National Center outlooks as guidance in decision support for partners. As the threat approaches, NWS forecasters provide more frequent and detailed assessments of possible hazards and impacts to partners through webinars, phone calls, briefing packages, and social media. Important to their role as local predictive experts is building relationships with core partners, like emergency managers, who use NWS information to make decisions themselves or help others to do so.

Emergency management across the United States is a formalized field of practice committed to enhancing the overall safety of citizens and preventing or reducing the impacts brought on by any number of environmental, technical, or human-made hazards. Across the country, emergency managers provide protection and watchful care according to specific jurisdictional boundaries across both public and private sectors. To that end, practitioners in the field operate at municipal, county, state, and/or federal levels, and they do so from physical spaces that span across government buildings, public health care facilities, educational institutions, or private businesses to name several. When severe weather threatens, EMs in the CWA studied access information produced by the local NWS WFO multiple times a day across various digital platforms to understand the likelihood and potential severity of conditions within their jurisdiction. With the local WFO's assessment of the weather hazard in hand, EMs then carry out a number of activities, such as communicating the potential weather threat onward to local representatives, deploying resources and assets to sites where potential hazard is expected to strike, and activating their Emergency Operations Center, to name a few.

Methods

In the winter of 2020, we conducted a multisited ethnography of a cool-season severe weather event, simultaneously deploying to three interrelated spaces that are central to tornado prediction and information dissemination for public safety in the United States. We chose this method for its ability to provide a new framing of confidence as one arising not within a single locus of prediction but across and between spaces as they interact, allowing us to "trac[e] and describ[e] the connections and relationships among sites [often] thought incommensurate" (Hine 2007). Further, this approach helped us understand confidence from particular standpoints, or the ability to see from the point of view of the other (Harding 2013), which generates greater translatability or commensurability. While the areas of responsibility for each of these three sites differed in their geographic extent (e.g., the CONUS versus a county) there was overlap in the area they each were assessing (e.g., where storms occurred) and communicating for.

To help us align our data collection processes, the authors held an intensive 2-day meeting months prior to deployment to practice, discuss, and refine our methodological approaches so we could identify common aspects of confidence at each site. For instance, we discussed what types of things constitute data to capture (e.g., office layout, interactions between and among people, information attended to) and how to write effective field notes. While in the field, we spent a few hours at the end of each day discussing features of confidence that emerged in daily observations and initial insights gleaned at each site, thus sensitizing our observations across sites the next day. Our goal was to employ a similar but not identical approach. Rather, we followed common ethnographic protocols for allowing our respective disciplinary perspectives to yield differences in our respective sites, as we illustrate below, and to enhance analysis and findings.

To identify analytic partners for this study, we followed approved Institutional Review Board protocols and contacted staff at the National Center and an NWS WFO located in the southeastern United States, based on relationships from previous research with these groups. We emailed key personnel to request permission from individual forecasters and NWS union officials to conduct observations at these sites. We then worked with the local NWS WFO management to identify interested partners in emergency management within their CWA. To recruit participants, emails were sent to directors of several OEMs, and one county-level office agreed to participate.

Sampling strategy.

Selection of tornado threat type. Working with our analytic partners, we held meetings 6 months before planned deployment to determine which types of severe weather observations were most useful to them. NWS forecasters noted their challenges in communicating marginal tornado threats given the lower predictability for these events, as well as possible widespread impacts not only of tornadoes but also wind and hail, which can be destructive for communities. EM staff agreed, noting additional challenges of communicating tornado risk when members of the public may not be as aware of cool-season tornado frequency. Based on these conversations, we selected two events to observe, one during the cool season (i.e., December–February) and one during the warm season (i.e., June–August). We deployed to the field and collected data for a cool-season event, further described below, but our plans to deploy for a warm-season event were disrupted in spring 2020 by COVID-19 restrictions on travel and access to federal and state offices.

The team agreed that a marginal cool-season tornado event would be selected based on a Storm Prediction Center (SPC) outlook for Marginal or Slight risks that covered areas of both the WFO's CWA and OEM. Importantly, because of our interest in how confidence emerges in these spaces before and during a possible event, our goal was to observe office activities and

communications in the days leading up to the event, regardless of whether or not the event verified. Nevertheless, it was challenging to deploy with sufficient lead time for observations given the uncertainties of forecasting a low-predictability, high-impact tornado threat. Coordination of schedules and travel among researchers and office staff also resulted in delayed attempts to deploy in the first year. Cancelled deployments also occurred due to a death in one of the analytic partner's communities, last-minute agency visitors to an analytic partner office, and illness among the team. Much like intense observing periods (IOPs) for physical science teams, multisited ethnographies provide essential observations but are carried out under complex data collection contexts.

Selection of tornado event deployment. To select a tornado event for observation, NWS forecasters notified the team via email when there was a chance for tornadoes in the areas of WFO and EM partners during the seasons of interest. We then held brief meetings to assess the feasibility of deployment, given the need to observe activities and communication leading up to and during the threat. This assistance from our analytic partners was essential since we could not spend weeks in the field waiting to observe multiple events, nor could we be in place to observe the earliest beginnings of the assessment of confidence for a given event.

In winter 2020, a day 6 SPC outlook with a 15% probability was issued for the area of interest. According to the SPC, this "highlighted area [is] equivalent to 2-SLGT-yellow"[1] risks (National Weather Service 2022), thus meeting our criteria. We monitored the situation, checking in with our analytic partners the next afternoon. Based on a continued 15% at day 5, we made the decision to deploy the next day, day 4, which put us each at our respective sites the morning of the SPC day 3 outlook, 2 days before the severe weather threat was predicted to occur.[2] While atmospheric signals were in place for a tornadic threat, there was still significant uncertainty about if and how the threat would materialize.

[1] SLGT is shorthand for slight risk.

[2] The day we arrived at our respective field sites also was a day 2 outlook for a risk of severe weather for areas that include our WFO CWA and OEM. This means there was a risk of severe weather the following day, which in turn means that our WFO CWA and OEM areas of study had a risk of severe weather 2 days in a row. The day 2 outlook for these areas was a Slight risk early that morning but was downgraded to a Marginal risk midday. Our analysis presented here does not directly focus on this threat, but it is important to acknowledge it as context for our data collection.

Data collection. Data collected at each site included audio recordings, field notes, forecast products issued by those we observed (e.g., outlooks, watches, warnings), as well as other messages exchanged and disseminated (e.g., email correspondence). We focused on capturing informal interviews and emergent conversations with participants in their respective offices, as well as interactions between staff as they worked. Our observations captured what data and other information were being attended to and evaluated (e.g., numerical weather prediction guidance, weather observations) by the different expert groups. Data also included recorded social science team conference calls held at the end of each day, during which we shared emerging insights about our respective sites that sensitized us to specific aspects of confidence in future observations. We left the field the day after the tornado threat occurred. In total, we collectively spent 75 h at three sites.

Observations were audio recorded for a total of 54 h of data. Because of the technical and complex nature of some recordings (e.g., multiple people talking together), the authors transcribed all material themselves over the course of several months, providing quality control of transcripts and incorporating field notes, photographs, and other relevant data. Material was deidentified and discussed iteratively among the authors during monthly meetings.

Analysis approach. The authors followed common qualitative analysis methods and independently coded the data in NVivo software, both deductively, examining instances of

confidence mentioned explicitly by participants, and inductively, identifying less explicit examples of the concept that emerged from different sources (e.g., webinars), expressions (e.g., tone of voice), representations (e.g., images), and meanings (e.g., collective or individual). Given the authors' different disciplinary backgrounds and their different sites of study, this analysis involved discussions of examples of confidence that illustrated various manifestations and contexts of the concept at each site. This process helped establish consistency in our analytic approach to the data while also facilitating knowledge transfer and codevelopment of new ideas among ourselves given our different areas of expertise.

Twice during the analysis process, we held virtual meetings with analytic partners to check for resonance, which is an important indicator of validity in ethnographic work (Tracy 2010), and for accuracy in understanding the content and context of material gathered. Finally, a copy of this manuscript was shared with analytic partners before submission to gather input, clarifications of information conveyed, and identify data that were incomplete or too sensitive. Given our ethical commitments to our ongoing relationships with analytic partners, including them in the publication process was crucial to the robustness and validity of analysis and findings.

Results

In this section, we describe confidence at each field site from three different disciplinary frameworks: risk communication, science and technology studies, and sociology. Each set of results below relate in different ways to the same tornado threat predicted to occur later in the week as designated in each section by day 3, or 3 days before a weather event; day 2, or 2 days before; or day 1, the day of a weather event. Within each site, multiple types of confidence occurred, but to parse the different manifestations of the concept here and to illustrate their differences and similarities, we here present only some of the ways confidence was illustrated by our data. Thus, any individual explanation of confidence is not meant to be exclusive of that specific site nor do individual examples reflect the complexity of confidence at various temporal and spatial scales. Finally, by selecting these disciplinary frameworks, we do not suggest they fully addresses the nuances of confidence; rather they reflect our collective disciplinary knowledge and the possibilities of explanations. For each site below, we first offer a brief summary of relevant theory from our respective disciplinary perspectives (in "Confidence in risk assessment and risk communication literature," "The social science of expectations," and "Culture of confidence: Culture, social institutions, and interaction ritual" sections). We immediately follow each disciplinary framing with the results from our respective field notes and transcribed recordings.

Site 1: The National Center and risk assessment.

Confidence in risk assessment and risk communication literature. In the field of risk assessment and risk communication, the concept of confidence emerges in multiple ways. For instance, in the trust, confidence, and cooperation model, confidence is deemed "instrumental and calculative" and defined as "the belief, based on experience or evidence (e.g., past performance), that certain future events will occur as expected" (Earle 2009). Interestingly, this work characterizes trust as "social and relational"—with the implication being that confidence is not. Another way that confidence emerges is in how experts subjectively judge the probability of a risk (Morgan 2014). Such expert elicitations are commonly used in the context of climate change to derive quantitative estimates of the likelihood of a risk. However, such quantitative estimates can be incorrect—that is, they can yield overconfident estimates when there is a true value known to be compared against (Morgan 2014). By comparison, for the fifth climate change assessment report, experts determined confidence based on two dimensions: 1) the type and amount of scientific evidence for something and 2) the

agreement among that evidence (Mach et al. 2014). These two factors are combined into an ordinal, qualitative judgment of confidence about the validity of findings (Fig. 3).

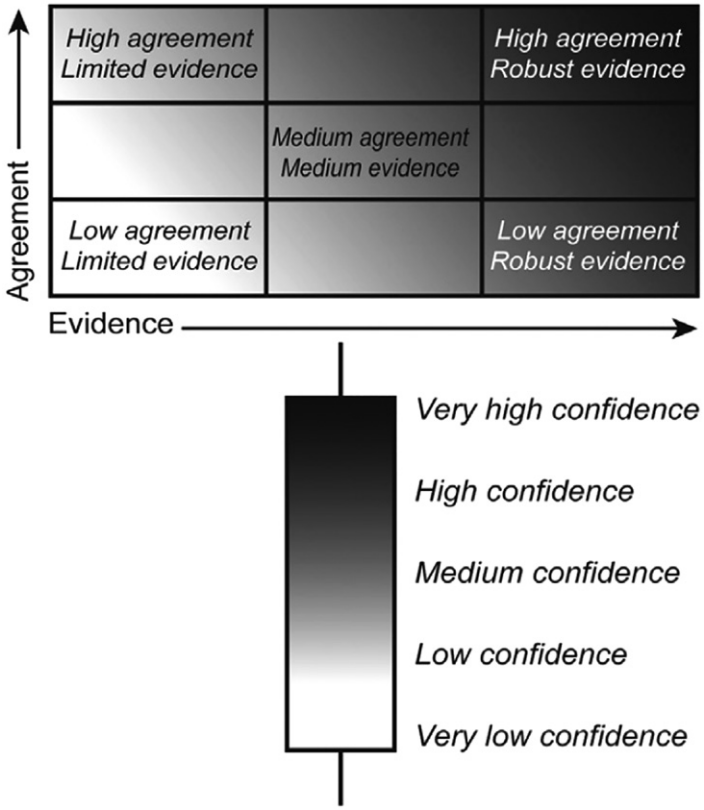


Fig. 3. IPCC AR5 expert judgment approach for evaluating confidence in climate change assessment findings (from Fig. 1 of Mach et al. 2014).

The ideas of confidence being based on experience or past performance coupled with it being a function of the amount of evidence and agreement among the evidence are all relevant to how National Center forecasters evaluate confidence when there is a threat of severe storms. Yet there are critical contextual differences. IPCC assessments involve hundreds of experts who spend months evaluating information and making risk assessments, and the quality of their judgments may not be able to be verified for years to come. Comparatively, NWS forecasters, including those at a National Center, make forecast decisions about severe storm risks individually or in small groups. Furthermore, they have new observations, statistical guidance, and numerical weather prediction (NWP) guidance streaming in on time scales of minutes to hours, and they must evaluate this constantly updating evidence under time pressures to issue both standard products (e.g., convective outlooks) by certain times as required by NWS policy directives and event-driven products (e.g., watches) with sufficient lead time for downstream users to receive and respond to the information. Forecasters also have the opportunity to verify the outcome of their predictions within minutes to days after issuing them, allowing forecasters to assess performance, including about model biases and other predictive uncertainties.

These contextual differences nuance how confidence emerges and operates for forecasters, and confidence can be even further complicated based on how it emerges in action for National Center forecasters' high-stakes task of issuing a convective outlook for a complex risk scenario.

National Center forecasters determining whether to upgrade a convective outlook. National Center forecasters' decision-making about what outlook category to issue begins with their meteorological expertise about what atmospheric ingredients must be in place and what processes must unfold for severe weather to occur. Such expertise also includes knowing what ingredients and processes can interfere with and reduce the threat of severe weather. Underpinning their meteorological expertise is their climatological expertise about the likelihood of atmospheric ingredients and processes occurring at different times of the year. Grounded in their severe weather expertise, forecasters interrogate many sources and types of data to determine the chance that severe weather hazards will occur. They look at observational data (e.g., from satellites, radar, soundings, surface stations) at the current time and in the recent past. They also look extensively at deterministic and ensemble-based NWP

guidance from coarse and convection-allowing models. Forecasters compare the guidance, looking for (in)consistencies that are critical to the possibility of severe weather. Forecasters also compare model predictions valid at current or recent past times against atmospheric observations to help them assess model errors in the occurrence, magnitude, and placement of different weather parameters.

During our cool-season event of study, it was day 2, the day before a severe weather threat, and National Center forecasters were evaluating the chance of tornadoes, severe wind, and severe hail in order to decide what convective outlook to issue. A Slight risk was already in effect, issued for the day 3 forecast. Thus, the specific decision the forecaster who was working the day 2 forecast desk faced was whether to upgrade the Slight to an Enhanced risk. As shown in Fig. 4, the Enhanced category is associated with probabilities of tornadoes, wind, and hail that are double the probabilities associated with a Slight risk. The decision therefore is not trivial.

The day 2 outlook must be issued by 1730 UTC, which was 1130 local time for all of the sites of study. The National Center forecaster had spent a couple of hours loading and analyzing myriad pieces of information, both in their National Center Advanced Weather Interactive Processing System (NAWIPS) workstation and from websites. They interrogated observations over multiple areas and at different layers of the atmosphere; multiple environmental and storm-specific parameters and thresholds from different models valid at many time steps; prior SPC forecasts; and questions and input from WFO forecasters via phone and NWS's internal chat system. As the outlook issuance time approached, the National Center

Day 2 Outlook Probability	TORN	WIND	HAIL
2%	MRGL	Not Used	Not Used
5%	SLGT	MRGL	MRGL
10%	ENH	Not Used	Not Used
10% with Significant Severe	ENH	Not Used	Not Used
15%	ENH	SLGT	SLGT
15% with Significant Severe	MDT	SLGT	SLGT
30%	MDT	ENH	ENH
30% with Significant Severe	HIGH	ENH	ENH
45%	HIGH	ENH	ENH
45% with Significant Severe	HIGH	MDT	MDT
60%	HIGH	MDT	MDT
60% with Significant Severe	HIGH	HIGH	MDT

Fig. 4. The categorical system and associated probabilities for tornadoes, severe wind, and severe hail for SPC's day 2 convective outlook (from www.spc.noaa.gov/misc/about.html).

forecaster conveyed the crux of the information they were evaluating and the consequent uncertainties.

> If we get 1500 [J/kg] MLCAPE [mixed-layer convective available potential energy] with 50, 60, 70 knots of effective shear—I mean it doesn't take a genius to figure out you got problems with convection developing. (...) My concern—this is surface-based CAPE from the NAM [North American Model]. It's got somewhere around 500 to 1000 [J/kg], maybe. I'm just not sure we're going to get 1500 [J/kg]. If I thought we were going to get 1500 [J/kg] MLCAPE tomorrow, I would say we need to go Enhanced. (...) Two main things that are holding me back. This widespread convection that forms with that northward advancing marine warm front. And it's not just the Euro [model]. You look at, for example, the GFS [Global Forecast System model] at 15, 18, 21 [Z]. There's a lot down here. (...) This is not an uncommon occurrence. As the low-level jet strengthens through the day, and you get the low-level mass response, and then you get just a rash of warm sector, basically low-level warm advection precip, and then along and north of that tends to get hosed. It just doesn't destabilize as much. (...) I'm using the 0Z run of the HRRR [High Resolution Rapid Refresh] as well. It's got some higher-end UH [updraft helicity] tracks right in there. The shear—that's effective bulk shear—50, 60 kts. That's incredibly good. That's common though. It's the instability that's holding me back. If I had any confidence [bangs his fist on the desk] that we would see more than 1000 [J/kg] MLCAPE—that's kind of my threshold for southeast severe. It's not a hard threshold, but if I think it's more likely than not that we're getting 1250 to 1500 [J/kg] with this amount of shear, then yeah, we probably need an Enhanced.

This quote from the National Center forecaster of a much more extensive analysis they were doing reveals the complex multitude of evidence being evaluated, the agreement and lack thereof among it, and how past performance of model guidance all come together to shape the forecaster's confidence. Specifically, they are particularly evaluating the magnitude of meteorological ingredients that are critical for severe weather occurrence, inconsistencies among the model guidance in whether the ingredients will reach those magnitudes, potential model error, evidence of inhibiting ingredients, and past model error, all of which result in low confidence about the probability that sufficient conditions will occur for severe weather the next day.

In addition to commenting about the evidence they were interrogating, the National Center forecaster also remarked about their thinking regarding the Enhanced category.

> I'm just not willing to commit to an Enhanced on a day 2 update because, basically, once it goes in, it's not coming out. We're stuck with it. And, in my personal opinion, I'm okay with Slights busting because it's a Slight. There's a lot of uncertainty. And Enhanced is category three out of five. (...) You get uncomfortable once you start going higher in those categorical outlooks. When you get to an Enhanced, it's really saying that we think this is more likely than not going to happen. And I'm not convinced of that yet.

To provide important context about this forecaster's quote, the National Center is not literally "stuck with" the upgraded outlook. Rather, they aim to minimize vacillations in the outlook categories from the issuance of one outlook to the next—especially if the change involves a downgrade—given how such messaging could adversely affect perceptions and responses to the risk. In addition, the forecaster saying they are "okay with Slights busting because it's a Slight" is commensurate with the low probability of 5% associated with a Slight risk as compared to the double 10% probability associated with an enhanced risk. These comments by the forecaster further reflect their lack of confidence in the chance that severe weather will occur. But, importantly, they also hint at elements of confidence that go beyond evidence to relational aspects regarding how users of the outlook perceive and make decisions based

on the product—and changes in the product from one forecast issuance to the next—and the corresponding implications of it not verifying. Accordingly, the National Center forecaster opted not to upgrade the day 2 outlook to an Enhanced risk.

Site 2: NWS WFO confidence through the lens of expectations.

The social science of expectations. Scholars in the social sciences who study the concept of expectation examine current circumstances, social and political dynamics, and ethical debates that shape how we attempt to create different types of possible futures. For example, our beliefs about a particular vision of equitable access might drive technological innovations and their policy implications (Borup et al. 2006). Individuals take actions that they hope will produce a desired future, sometimes called a sociotechnical imaginary or speculative future (Jasanoff 2015). This involves not only tangible societal outcomes but anticipation about how people will behave or which choices they will make, as in, for example, individual economic investments (Dequech 1999). Expectations, then, are not neutral. They have the power to mobilize resources, fund infrastructure, and shape society in planned and unforeseen ways. Over time, this anticipation includes "the symbols, practices, tools, and patterned ways" that a given "social world answers the questions 'what is next?' and 'now what?'" (Baker 2021). The present and the future are tied together through a set of beliefs, experiences, and practices that inform our expectations. Such social and cultural cues likewise provide opportunities for understanding what will or should happen in the future.

Important to answering questions about what will happen next is the notion of confidence and how it mitigates those uncertainties that influence decisions made in the present moment. In one view of this relationship between confidence and uncertainty, confidence is "jointly determined by uncertainty perception and uncertainty aversion" (Dequech 1999). Uncertainty, in this case, reflects both knowledge of the world (e.g., perception) and how we are oriented toward it, or our disposition (see Keynes 1937). In other views, expectation can provide an informational source for confidence, as when lessons learned about one's own decision-making from past experiences allows one to predict more accurately outcomes of future decisions (Boldt et al. 2019).

Yet, as is mentioned in the previous section on confidence at a National Center, knowledge is not always an individual endeavor or discrete outcome; it materializes within the context of others and can come to represent rituals that influence a shared confidence, as discussed in the section, "Site 3: EMs and ritualized interaction as shaping a culture of confidence." As this section will show, confidence may similarly emerge as a function of relationships built over time. Within the NWS WFO, confidence shapes the uncertainties relative to the temporal unfolding of behaviors regarding both how people might or should behave and what the atmosphere will do.

The analysis below examines excerpts from a webinar created on day 3 and illustrates how expectation influences two types of confidence. The first derives from forecasters' knowledge about the weather itself, or what might be thought of as a confidence of expertise. It highlights how skill gained from education, scientific tools, and past weather events generate beliefs about how the weather will play out. The second type is based on previous experiences with partners in emergency management, or what might be thought of as a relational confidence. It illustrates how NWS forecasters' expectations about partners' future actions are present in their communications.

Expectations as confidence in expertise about the weather. On day 3, 2 days prior to the potential severe weather event, NWS WFO staff began preparing partners in public safety for possible weather outcomes. Preparations included hosting a webinar briefing highlighting forecasters' expectations for the likelihood, timing for, severity, and impacts of tornadoes

and flooding in their CWA. According to field notes, staff have an informal policy stating that if severe weather were to increase in likelihood or severity, as with an SPC Enhanced risk, another webinar briefing might be added on day 2 to give partners adequate time to make preparations based on the update. For this event, it was unclear on day 3 whether or not the Slight risk would be elevated to an Enhanced risk within the next 24 h.

For this briefing, one forecaster created a PowerPoint presentation in consultation with other staff in the office, the final version of which was meant to reflect WFO staff beliefs about hazardous weather that might occur throughout the week. Webinars were recorded live with an audience of emergency managers on a conference call who could view slides as the forecaster talked through them. At the end, the audience was given an opportunity to ask questions; the final video was posted on the office YouTube channel.

After highlighting details about the synoptic setup for their region, the forecaster explained the potential for tornadoes in their CWA for the next 48 h:

> Now by [day 1] that cold front should have moved through into the state. And with that boundary, we're going to still see some storms ahead of the boundary and along it and then towards [the] evening, that increasing shear, as well as a little bit of instability. It's put us in a [SPC] Marginal risk. Again, damaging winds is going to be our primary concern, but tornadoes cannot be ruled out given the amount of shear that is setting up ahead of boundary. Now, the timing for that is [day 1, afternoon and evening]. Prior to this, we looked a little bit wider of an area of concern for central [state name]. But with the last few runs of the models, the concern for the higher risk for severe weather has actually shifted to the coast. Now, there is a lot of uncertainty as to where that frontal boundary will actually set up, so if that frontal boundary moves a little bit further back, we could see some of our Marginal risk shifting.

Expectations play out in two interrelated ways in this excerpt: First, forecast details in this webinar were couched in the language of plural pronouns of "we" and "our," representing a common voice of expertise generated through complex processes of scientific discussion of the weather, as observed in field notes. These included individual and collaborative assessment of data (e.g., computer models and real-time observations), interoffice conversations about meteorological principles, and examination of analogous weather scenarios. For example, the webinar included referenced changes in updated runs of computer models, which helped convey the uncertainty about the severe weather's location, timing, and scope. Noting the SPC's Marginal risk similarly offered additional evidence to bolster partner confidence in forecasters' predictions and caveats about potential severe weather in the coming days. Choices about which content to include and explanations of what that content means may reinforce partners' confidence in the forecasters' predictions, creating a kind of interactive confidence among expert groups. Within the context of a public briefing, the content of the slides reflected the forecasters' collective expert confidence in their expectations for future weather.

Expectations about partner actions as a relational confidence. A second type of confidence born out of expectation is further evidenced in the briefing, one reflected in the desired actions forecasters wanted partners to take. Publicly communicating their expert confidence generated certain expectations of the audience of emergency managers. At the conclusion of the webinar, the forecaster offered a summary of what partners would need to do to be prepared for the evolving weather:

> So, at this point, what we're making sure you know is that this forecast could very easily change, and you're going to need to be paying attention for the next couple of days prior to [the afternoon of day 1], as far as our [tornado] threat is concerned.

There is an expectation that emergency managers will act vigilantly and check back with forecasters. After the webinar ended, when asked about this closing statement to partners, the forecaster elaborated on how they hoped partners would behave and why.

> I want EMs to be aware that this [forecast] is not a slam dunk so that they don't lose situational awareness.... So I am far more likely to say, "Hey, guys, look, this is what it says now. Don't stop watching." Yeah, that's gonna be my bigger punch to them: Don't stop watching. So you give them enough uncertainty that they go, "Hm, I need to keep watching." Because if you just said, this is what's going to happen, this is what's gonna happen, this is what's gonna happen, they could tune out because they're not going to get another briefing before [day 1] unless this [tornado threat] really ramps up.

The directive to partners to continue checking the forecast is itself an action predicated on a desired future, one in which EMs attend to the weather at the suggestion of WFO staff. Confidence in this example arises from the expectation that partners who continually "watch" the weather will benefit from the forecaster's expertise about this uncertainty as it evolves. As discussed above, to motivate this watchfulness, forecasters provided illustrations of their expertise in the slides and narration of the webinar, representing the staff's collective expectations about the likelihoods, timing, and locations of threats. Forecaster confidence about weather prediction, then, is intimately bound both to the nuances of risk information itself and to what partners may do with this information.

From observations made of office practices, forecasters based their knowledge of partners' need for different types of information and the timing of its delivery on past discussions with them after severe weather events, as well as through partner requests for updated risk information through phone calls, texts, and emails in the immediate hours leading up to and during severe weather. Past experience shapes a relationship-based confidence that informs advice to partners on what to do next. One challenge for forecasters is that their confidence about partner behavior may be premised on a singular imaginary (Jasanoff 2015) of the activities and responsibilities that shape partners' individual decisions, which may or may not hold true in the future. For instance, the data above illustrate that the forecaster believes EMs will return to them to get updated forecast information given the uncertainty, and this belief is based on who they know and thus what they imagine their EMs will do. But there is a possibility that an EM might instead seek a forecast from someone else who can give them more certain information. Thus, the complexities of partner decisions require forecasters to continually hone their own situational awareness about the decision spaces and expectations of their partners. The relational nature of IDSS, then, is central to developing mutual confidence, or as discussed in the literature of risk assessment above, of trust between forecasters and their partners.

As the next section will show, confidence may also emerge from the institutional rules and patterned interactions to create a collective confidence that becomes visible when disrupted.

Site 3: EMs and ritualized interaction as shaping a culture of confidence. Emergency managers and staff in the OEM studied here have accumulated over 50 years of collective experience representing their jurisdiction and have worked together as a team for many years. This has fostered a cohesive relational dynamic among the group, which was evident during our cool-season observation when members of the team were heard finishing each other's sentences, and where members of the team were seen to move in and out of multiple, ongoing conversations with relative ease. Like offices around the nation, this OEM is responsible for a wide range of emergencies, both weather and nonweather related, and so must meet the expectations of a variety of stakeholders and publics.

Located in a larger metropolitan area, this OEM is higher resourced than others in more rural areas, which shapes staff's ability to respond to severe weather threats and their robust relationship to the local WFO. For example, the director of this OEM has the capacity to choose from and allocate a range of assets depending on the group's assessment of needs, and they have sophisticated systems in place for communicating and responding to these types of threats. Despite having extensive practical knowledge regarding impacts brought on by severe weather events and a basic comprehension for when and how tornadoes occur—they report that in their area tornadoes tend to happen mostly at night and when there is "a battling between warm and cold fronts," for example—this experiential knowledge does little to support their confidence regarding when, where, and how intensely tornadoes will strike specifically. For this, EMs in this study give deference to meteorologists' actions at the local WFO.

In this section, we draw on the notion of deference and suggest that it has come to represent, and be representative of, a culture of confidence between EMs and the local WFO. To accomplish this, we draw on the concept of social institutions (Tansey and Rayner 2008) and Goffman's (1967) theory of interaction ritual, macro- and microinteractional approaches to sense making and action, respectively. Our purpose is to show how broader social institutions play a foundational role in setting the stage for social conduct between EMs and meteorologists and, and when coupled with microlevel ritualized interaction, how an expectation and ordering of behavior is established. From this, a culture of confidence is inscribed.

Culture of confidence: Culture, social institutions, and interaction ritual. Culture is an integral part of the systems that shape social life. It is something that is practiced, adopted, and performed within and across social groups. Culture operates under a distinct set of rules that invoke specific practices in various social settings, may that be in homes, workplaces, or organizations, to name a few. At the same time, as Fine and Harrington (2004) report, culture is not simply a theoretical variable or a fixed category; it is adaptable and flexible. Further, these scholars assert that people *do things* with culture, such as make claims, deploy it in negotiations of everyday life, use it to create affiliation, and sometimes even to shape civic life. A number of elements contribute to the establishment of cultures. Here we focus on social institutions and interaction rituals among and across groups.

Social institutions refer to the structural mechanism by which people are organized in society (Tansey and Rayner 2008). They can be thought of as large-scale forces that help to shape and pattern behavior. Likewise, they help to establish and maintain social order. These organized patterns of behavior come to be represented by social norms and conventions that guide interactions within and across groups. While social institutions set the stage for interaction at the macro level, interaction ritual operates locally to establish a culture of confidence. Interaction ritual as a theoretical concept originates in Emile Durkheim's (1912) analysis of the social basis of religion but was extended by Goffman (1967) and Collins (2000) to show its relevance for explaining the importance of mini rituals in setting expectations in everyday social life. In keeping with Collins (1987, 2004), interaction rituals are procedures between two or more people under conditions of copresence. They comprise negotiations that represent moments of shared social reality, require cooperation and a mutual focus of attention, and are dependent upon shared motivations and resources. Interaction rituals set typical expectations of conduct by members within and across groups to which people adhere or transgress.

Transgression of the typical expectations is neither good nor bad but a disruption of an established pattern and significant in the way that they illustrate the presence/importance of the ritual. At the same time, transgressions in the interaction ritual have the potential to generate unintended consequences.

Culture of Confidence between EMs and the local WFO: Interaction ritual and an example of transgression. Through the interaction ritual, unique operational rhythms of interaction exist between EMs and forecasters' actions. For example, leading up to and during severe weather, EMs in the CWA studied have come to expect a certain sequence of communication by the local WFO. This communication sequence is perceived to be carried out by the local forecasters with a great deal of consistency and regularity from one event to the next and includes: a weekly morning briefing with other partners, updates issued via social media throughout the week with changes in severe weather risk level from SPC, a midweek conference call, communication of a watch with approximately 3 h of lead time, and finally, a warning issued shortly in advance of impacts. When communication from the local WFO—whether originated by them or passed along by them—follows this structure, forecasters' interpretations and actions surrounding tornado risk become more recognizable to EMs. In other words, local WFO communication leading up to and during severe weather is akin to a mini ritual that has, over time, established confidence for EMs regarding the types, timing, and extent of communication provided by the local WFO in severe weather conditions.

To illustrate this ritual, consider the EMs' response to a tornado watch issued by SPC for their jurisdiction and communicated by the local WFO on day 1, the day of the predicted severe weather events. A group of four representatives from the EM office studied and a coauthor are having lunch at a local restaurant when the alert is issued. From field notes:

> *EM2's phone bings*. He says to the group, "Under a tornado watch." EM1 and EM3 take a break from eating and pick up their phones. EM1 repeats, "Just out. Under a tornado watch." EM3 is concerned about the local WFO's Facebook page, "But they haven't updated their [Facebook] since this morning," the coauthor reports. EM2 responds, "I got it on my radar app. See tornado watch for [the EMs' jurisdiction]. Expires 6 hours and 13 minutes …" EM1 tells the group, "[this] means we'll transition [activate the Emergency Operation Center] when we get back to the [office]."
>
> We each take time to finish the food on our plates, then take turns paying for the meals (five in total). The mood is light, the EMs are calm; there is no rushing and it seems we return to the office not with a great sense of urgency but patiently and with purpose.

In this scenario, there was no official warning issued for the OEM's jurisdiction that afternoon; however, there were 3 h 13 min of lead time between the tornado watch and the first observation of heavy rains in the southeastern part of their area. This is in keeping with the "approximately 3 hours of lead time" mini ritual that has been established between these two social groups. The calm response of the EMs at lunch during this observation suggests that forecaster conduct was in accordance with the interaction ritual, and from this, the culture of confidence was maintained.

As an illustration of the opposite, there have been circumstances where forecasters' communication has occurred in unexpected ways. When this happens, the established culture of confidence between EMs and the local forecasters is shaken, or as Collins (2004) suggested, there was a lapse in shared social realities. Consider the following example from 2019, when the local WFO issued a warning for an area that included this OEM under study without a watch being issued beforehand by SPC:

> EM1: There's another event that we had a tornado warning
>
> EM3: Came out before [the watch]

EM1: [Before] a watch. We were all at home asleep.

EM3: That was terrible.

EM1: [...] We had a tornado warning with no watch, which means we wasn't here [at the office]. So I don't like getting caught like that. It happens.

EM3: Not very often.

The example above represents a transgression in the interaction ritual, represented by the presence of divergent social realities: one group did not issue a tornado watch given their assessment of atmospheric conditions, whereas the mini ritual established between the local WFO and the OEM informed the EM's expectation that prior to a tornado warning a tornado watch would be communicated. This divergence brought on a shakiness in the culture of confidence, one that manifested as EMs "getting caught" unaware, something that was considered "terrible" by EM3 and something that is "not liked," according to EM1. This example demonstrates how disorderly and disruptive ruptures in the interaction ritual can be. The jump straight to warning with no watch created "chaos" according to EM1, "because we're not here [at the office]," which made it "a pretty hectic day." Transgressions of interaction rituals are meaningful to capture since they expose the importance of the established rhythm of activities for the way EMs make sense of the situation, including their subsequent behaviors and actions. For this particular group of EMs, transgressions of the prevailing pattern confirm how this group of practitioners have come to rely on interaction rituals to shape their expected sequencing of forecaster communication, which for them helps coordinate a calm and orderly response.

Overall, observations with EMs revealed how interaction rituals help construct a culture of confidence between their group and the WFO, which enabled EMs to *do things* with this culture, such as communicate or respond to threats. Interestingly, the "warning with no watch" transgression also reveals that the culture of confidence for these EMs is not just based on their interaction with the local WFO. Rather, unbeknownst to the EMs, activities of a third entity external to the immediate interaction—in this case, the agency responsible for issuing tornado watches—contributes to a disruption in the interaction ritual and a shaken culture of confidence.

Discussion and conclusions

We conducted a multisited ethnography with analytic partners in an NWS National Center, an NWS WFO, and a local OEM to understand how confidence emerges within expert groups and propagates between them in the days leading up to a cool-season tornado threat. Focusing on a cool-season tornado threat in the southeastern United States reflected our analytic partner's interests in better understanding issues of predictability and impact on communication with partners and public. Our analysis was aimed at understanding the ways that confidence might differ in its manifestations and exhibit what Duguid (2005) calls a complex "social life." Three primary insights emerged.

First, we found that confidence is varied and complex in weather prediction, differing in how it is deployed, the scales at which it operates, what it might mean, and its function within the risk assessment and communication processes and sociocultural contexts. Specifically, we found that leading up to the cool-season tornado threat, NWS National Center forecasters assessed and compared a variety of evidence including observational data, deterministic and ensemble-based guidance, and input from others in their office and the local WFO. Confidence emerged in this process amid the context of their own expertise and their role as scientists

whose responsibility it is to accurately characterize tornado risk at a national level. During the same time frame, NWS WFO staff were both assessing the threat for multiple hazards, including the tornado threat, and synthesizing their predictions to convey likely risks to partners in public safety. While confidence within the WFO emerged through similar processes as the National Center given their common role as scientists, results focused on the communication of this confidence to partners in emergency management. Here confidence is a relational dimension of their work, one based on expectations and assumptions about the behavior of the weather and of emergency managers in light of uncertainty information. For emergency managers, for whom weather is only a small part of their daily responsibilities, they rely on local WFO staff to apprise them of weather risk through textual products, social media posts, and direct/indirect interactions leading up to a hazard. Confidence at this site reflected the establishment of patterns and rituals between themselves and NWS forecasters, creating a cultural norm that guided daily expectations for information and communication. In this case study, confidence operated through information but also through expertise and relationships.

As this analysis illustrates, confidence is not a concept that is deployed and understood in a singular way, nor is it static. Different types and manifestations of confidence exist, which are dynamic and difficult to anticipate. Confidence may cascade between groups, for instance, during webinars. Or it may be unexpectedly disrupted, as when risk information is presented in unfamiliar ways, like a warning that precedes a watch. When an expression of confidence is conveyed or received, then there is potential for mismatches, misunderstandings, and gaps between them, some of which may have important consequences, like a loss of trust between groups or misunderstandings that impact decisions. Confidence circulates as part of evolving situations and through varying types of relationships, which makes confidence a particularly challenging concept with which to grapple.

Second, while our results highlight specific dimensions of confidence at play in our respective sites, it is important to note that these examples are not mutually exclusive. Confidence generated by evidence and assessment, through expectations for others, and as a part of cultural norms is evident at all sites. It is also a feature of the varying roles (e.g., forecaster, public safety official) that different groups play in the weather communication process. Thus, confidence is multiple and complex, no matter the specific context within weather prediction. In part, this messiness is a function of the multifaceted aspects of weather prediction itself. It is likewise a result of the complicated social processes at play in the provision of high stakes information about weather risks and the ethical responsibilities experts have for the publics they serve.

Third, given that confidence is so critical to understanding weather risk and its many expressions, it is important for the weather community to attend more directly and deeply to other concepts that may be taken for granted as having a shared meaning, such as uncertainty, decision support, or effective communication. After all, how we understand and frame concepts that are crucial to weather prediction has tangible impacts on societal outcomes—these concepts become embedded in our technologies, influence our communication, guide policy implementation, and shape funding for research. In this work, we have demonstrated that confidence is alive in our social, political, and material worlds. In this effort, we follow others' efforts to take what is tacit and make it visible. For example, other key concepts, like vulnerability (Wisner and Luce 1993; Cutter 1996; Adger 2006) or disaster (Blaikie et al. 1994; Fothergill 1996; Oliver-Smith 1999; Quarantelli 1998), have been acknowledged by social scientists for decades as polysemous in meaning and in need of richer understandings.

One limitation of our study is the lack of ability to generalize from these three offices to others in the weather enterprise. Each WFO and OEM represents a specific subculture composed of experiences and expertise, informal policies and practices, as well as resources.

Case studies, however, offer important insights that can be carried out in similar spaces to deepen comparisons. Our data collection was limited to specific expert groups in the integrated warning team, excluding broadcast meteorologists and other public safety officials. Nor do we focus on the public. Given the nuanced ways confidence appears in the groups studied, it is reasonable to assume similar complexity exists for different publics and other experts. Future research about how members of these groups understand, articulate, receive, and attend to confidence dimensions of predictive information is needed.

Our goal has been to illustrate the importance of the concept of confidence and our need for a more robust interdisciplinary analysis and explication of it. We would do well to center these types of concepts in our scholarship to help characterize their many manifestations, but also how they live and travel among experts, between experts and the publics they serve, and within the public itself.

Acknowledgments. This research was funded by NOAA Grant NA18OAR4590311 for both Henderson and some of Spinney's time. NOAA Grant NA18OAR4590312 supported Demuth's time for this research. We wish to thank the NWS forecasters and emergency management officials who contributed their time to this project throughout our planning process, data collection and analysis, and feedback on paper drafts. Views and opinions in this document are those of the authors. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Data availability statement. Field notes, data recordings, and transcripts are archived and indexed on the authors' personal computers. Data availability to other researchers is in accordance with human subjects research and Institutional Review Board guidelines at University of Colorado Boulder.

References

Adger, W. N., 2006: Vulnerability. *Global Environ. Change*, **16**, 268–281, https://doi.org/10.1016/j.gloenvcha.2006.02.006.

Baker, Z., 2021: Anticipatory culture in the Bering Sea: Weather, climate, and temporal dissonance. *Wea. Climate Soc.*, **13**, 783–795, https://doi.org/10.1175/WCAS-D-21-0066.1.

Balazs, C. L., and R. Morello-Frosch, 2013: The three Rs: How community-based participatory research strengthens the rigor, relevance, and reach of science. *Environ. Justice*, **6**, 9–16, https://doi.org/10.1089/env.2012.0017.

Belone, L., and Coauthors, 2016: Community-based participatory research conceptual model: Community partner consultation and face validity. *Qual. Health Res.*, **26**, 117–135, https://doi.org/10.1177/1049732314557084.

Blaikie, P., T. Cannon, I. Davies, and B. Wisner, 1994: *At-Risk Vulnerability and Disasters.* Routledge, 496 pp.

Boldt, A., A. M. Schiffer, F. Waszak, and N. Yeung, 2019: Confidence predictions affect performance confidence and neural preparation in perceptual decision making. *Sci. Rep.*, **9**, 4031, https://doi.org/10.1038/s41598-019-40681-9.

Borup, M., N. Brown, K. Konrad, and H. Van Lente, 2006: The sociology of expectations in science and technology. *Technol. Anal. Strategic Manage.*, **18**, 285–298, https://doi.org/10.1080/09537320600777002.

Cavanaugh, D., M. Huffman, J. Dunn, and M. Fox, 2016: Connecting the dots: A communications model of the North Texas Integrated Warning Team during the 15 May 2013 tornado outbreak. *Wea. Climate Soc.*, **8**, 233–245, https://doi.org/10.1175/WCAS-D-15-0047.1.

Collins, R., 1987: Interaction ritual chains, power and property: The micro-macro connection as an empirically based theoretical problem. *The Micro-Macro Link*, J. C. Alexander et al., Eds., University of California Press, 193–206.

——, 2000: Reflexivity and social embeddedness in the history of ethical philosophies. *The Sociology of Philosophical Knowledge,* Springer, 155–178.

——, 2004: *Interaction Ritual Chains.* Princeton University Press, 464 pp.

Cutter, S. L., 1996: Vulnerability to environmental hazards. *Prog. Hum. Geogr.*, **20**, 529–539, https://doi.org/10.1177/030913259602000407.

Demuth, J. L., B. H. Morrow, and J. Lazo, 2009: Communicating weather forecast uncertainty: An exploratory study with broadcast meteorologists. *Fourth Symp. on Policy and Socio-Economic Research*, Denver, CO, Amer. Meteor. Soc., 3.1A, https://ams.confex.com/ams/89annual/techprogram/paper_152369.htm.

——, J. K. Lazo, and R. E. Morss, 2011: Exploring variations in people's sources, uses, and perceptions of weather forecasts. *Wea. Climate Soc.*, **3**, 177–192, https://doi.org/10.1175/2011WCAS1061.1.

Dequech, D., 1999: Expectations and confidence under uncertainty. *J. Post Keynesian Econ.*, **21**, 415–430, https://doi.org/10.1080/01603477.1999.11490205.

Doswell, C. A., III, A. R. Moller, and H. E. Brooks, 1999: Storm spotting and public awareness since the first tornado forecasts of 1948. *Wea. Forecasting*, **14**, 544–557, https://doi.org/10.1175/1520-0434(1999)014<0544:SSAPAS>2.0.CO;2.

Duguid, P., 2005: 'The art of knowing': Social and tacit dimensions of knowledge and the limits of the community of practice. *Inf. Soc.*, **21**, 109–118, https://doi.org/10.1080/01972240590925311.

Durkheim, E., 1912: The elementary forms of religious life. *Emile Durkheim: An Introduction to Four Major Works*, R. A. Jones, Ed., Sage Publications, 115–155.

Earle, T. C., 2009: Trust, confidence, and the 2008 global financial crisis. *Risk Anal.*, **29**, 785–792, https://doi.org/10.1111/j.1539-6924.2009.01230.x.

Fine, G. A., and B. Harrington, 2004: Tiny publics: Small groups and civil society. *Sociol. Theory*, **22**, 341–356, https://doi.org/10.1111/j.0735-2751.2004.00223.x.

Fischhoff, B., and D. MacGregor, 1982: Subjective confidence in forecasts. *J. Forecasting*, **1**, 155–172, https://doi.org/10.1002/for.3980010203.

Fothergill, A., 1996: Gender, risk, and disaster. *Int. J. Mass Emerg. Disasters*, **14**, 33–56, https://doi.org/10.1177/028072709601400103.

Goffman, E., 1967: *Interaction Ritual: Essays in Face-to-Face Behaviour.* Doubleday, 270 pp.

Harding, S., 2013: Rethinking standpoint epistemology: What is "strong objectivity"? *Feminist Epistemologies*, L. Alcoff and E. Potter, Eds., Routledge, 49–82.

Henderson, J., and M. Liboiron, 2019: Compromise and action: Tactics for doing ethical research in disaster zones. *Disaster Research and the Second Environmental Crisis*, J. Kendra, S. G. Knowles, and T. Wachtendorf, Eds., Springer, 295–318.

Hine, C., 2007: Multi-sited ethnography as a middle range methodology for contemporary STS. *Sci. Technol. Human Values*, **32**, 652–671, https://doi.org/10.1177/0162243907303598.

Hogan Carr, R., B. Montz, K. Maxfield, S. Hoekstra, K. Semmens, and E. Goldman, 2016: Effectively communicating risk and uncertainty to the public: Assessing the National Weather Service's flood forecast and warning tools. *Bull. Amer. Meteor. Soc.*, **97**, 1649–1665, https://doi.org/10.1175/BAMS-D-14-00248.1.

Jasanoff, S., 2015: Future imperfect: Science, technology, and the imaginations of modernity. *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*, S. Jasanoff and S. H. Kim, Eds., University of Chicago Press, 1–33.

Joslyn, S., and S. Savelli, 2010: Communicating forecast uncertainty: Public perception of weather forecast uncertainty. *Meteor. Appl.*, **17**, 180–195, https://doi.org/10.1002/met.190.

Keynes, J. M., 1937: The general theory of employment. *Quart. J. Econ.*, **51**, 209–223, https://doi.org/10.2307/1882087.

Kox, T., and A. H. Thieken, 2017: To act or not to act? Factors influencing the general public's decision about whether to take protective action against severe weather. *Wea. Climate Soc.*, **9**, 299–315, https://doi.org/10.1175/WCAS-D-15-0078.1.

Lazo, J. K., R. E. Morss, and J. L. Demuth, 2009: 300 billion served: Sources, perceptions, uses, and values of weather forecasts. *Bull. Amer. Meteor. Soc.*, **90**, 785–798, https://doi.org/10.1175/2008BAMS2604.1.

Mach, K. J., M. D. Mastrandrea, P. T. Freeman, and C. B. Field, 2014: Unleashing expert judgment in assessment. *Global Environ. Change*, **44**, 1–14, https://doi.org/10.1016/j.gloenvcha.2017.02.005.

Morgan, G. M., 2014: Use (and abuse) of expert elicitation in support of decision making for public policy. *Proc. Natl. Acad. Sci. USA*, **111**, 7176–7184, https://doi.org/10.1073/pnas.1319946111.

Morris, D., and Coauthors, 2008: The National Severe Weather Workshop scenario: Interactive adult learning for Integrated Warning Team partners. *17th Symp. on Education*, New Orleans, LA, Amer. Meteor. Soc., J3.6, https://ams.confex.com/ams/88Annual/techprogram/paper_131731.htm.

Murphy, A. H., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, **8**, 281–293, https://doi.org/10.1175/1520-0434(1993)008<0281:WIAGFA>2.0.CO;2.

National Weather Service, 2019: Impact-based decision support services. National Weather Service Policy Directive 10-24, 19 pp., www.nws.noaa.gov/directives/sym/pd01024curr.pdf.

——, 2022: SPC products. Accessed 22 May 2022, www.spc.noaa.gov/misc/about.html.

Oliver-Smith, A., 1999: "What is a disaster?": Anthropological perspectives on a persistent question. *The Angry Earth: Disasters in Anthropological Perspective*, A. Oliver-Smith and S. Hoffman, Eds., Routledge, 32–48.

Padilla, L. M., M. Powell, M. Kay, and J. Hullman, 2021: Uncertain about uncertainty: How qualitative expressions of forecaster confidence impact decision-making with uncertainty visualizations. *Front. Psychol.*, **11**, 579267, https://doi.org/10.3389/fpsyg.2020.579267.

Quarantelli, E. L., 1998: *What is a Disaster? Perspectives on the Question.* Routledge, 326 pp.

Su, C., J. N. Burgeno, and S. Joslyn, 2021: The effects of consistency among simultaneous forecasts on weather-related decisions. *Wea. Climate Soc.*, **13**, 3–10, https://doi.org/10.1175/WCAS-D-19-0089.1.

Tansey, J., and S. Rayner, 2008: Cultural theory and risk. *Handbook of Risk and Crisis Communication*, R. L. Heath and H. D. O'Hair, Eds., Routledge, 53–79.

Tracy, S. J., 2010: Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qual. Inq.*, **16**, 837–851, https://doi.org/10.1177/1077800410383121.

Williams, P., Jr., 1951: The use of confidence factors in forecasting. *Bull. Amer. Meteor. Soc.*, **32**, 279–281, https://doi.org/10.1175/1520-0477-32.8.279.

Wisner, B., and H. R. Luce, 1993: Disaster vulnerability: Scale, power and daily life. *GeoJournal*, **30**, 127–140, https://doi.org/10.1007/BF00808129.