Collaboration among Forecasters to Issue Severe Weather Hazard Information and Warnings Using the Hazard Services–Probabilistic Hazard Information (HS-PHI) Tool

JOSEPH JAMES,[a] CHEN LING,[a] ALYSSA BATES,[b,c] GREGORY J. STUMPF,[d,e] KIM KLOCKOW-MCCLAIN,[b,f] PAT HYLAND,[b,f] JIM LADUE,[c] KODI L. BERRY,[f] AND KEVIN MANROSS[d,g]

[a] *University of Akron, Akron, Ohio*
[b] *Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma*
[c] *NOAA/NWS/Warning Decision Training Division, Norman, Oklahoma*
[d] *Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado*
[e] *NOAA/NWS/Office of Science and Technology Integration/Meteorological Development Laboratory, Silver Spring, Maryland*
[f] *NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma*
[g] *NOAA/OAR Global Systems Laboratory, Boulder, Colorado*

(Manuscript received 4 August 2021, in final form 31 July 2022)

ABSTRACT: This project tested software capabilities and operational implications related to interoffice collaboration during NWS severe weather warning operations within a proposed paradigm, Forecasting A Continuum of Environmental Threats (FACETs). Current NWS policy of each forecast office issuing warnings for an exclusive area of responsibility may result in inconsistent messaging. In contrast, the FACETs paradigm, with object-based, moving probabilistic and deterministic hazard information, could provide seamless information across NWS County Warning Areas (CWAs). An experiment was conducted that allowed NWS forecasters to test new software that incorporates FACETs-based hazard information and potential concepts of operation to improve messaging consistency between adjacent WFOs. Experiment scenarios consisted of a variety of storm and office border interactions, fictional events requiring nowcasts, and directives that mimicked differing inter-WFO warning philosophies. Surveys and semi-structured interviews were conducted to gauge forecasters' confidence and workload levels, and to discuss potential solutions for interoffice collaboration and software issues. We found that forecasters were able to adapt quickly to the new software and concepts and were comfortable with collaborating with their neighboring WFO in warning operations. Although forecasters felt the software's collaboration tools enabled them to communicate in a timely manner, adding this collaboration increased their workload when compared to their workload during current warning operations.

KEYWORDS: Forecasting; Forecasting techniques; Numerical weather prediction/forecasting; Operational forecasting; Probabilistic Quantitative Precipitation Forecasting (PQPF)

1. Introduction

The current NWS severe weather warning paradigm can result in messaging inconsistencies along geopolitical boundaries. The NOAA NWS includes 122 Weather Forecast Offices (WFOs). Although these WFOs are tasked with issuing several short-fused, polygon-based warning products for their areas of responsibility [called County Warning Areas (CWA); Fig. 1], this study focuses on severe thunderstorm warnings (SVR) and tornado warnings (TOR). The NWS issues warnings as storm-based polygons, which are intended to represent the area that the hazard is expected to affect for the duration of the warning.

The Advanced Weather Interactive Processing System, second generation (AWIPS II) workstation is used by NWS forecasters to analyze and diagnose weather information and to issue forecasts and warnings. The current warning generation software (WarnGen) allows forecasters to define warning attributes, such as type (SVR and TOR), duration, threat, and others. Static 2D storm-based polygons define the warning areas. When polygons extend to CWA boundaries, the software automatically truncates the warning's area so that it only includes areas within the CWA of warning issuance (NWS 2020). Figure 2 illustrates a 3-yr grid of warning density that reveals distinct discontinuities between counties, and especially between CWAs (as seen by comparing both maps, with the CWA borders removed for one). If a warned storm (Fig. 3a) approaches and traverses a CWA border, a neighboring WFO assumes a shared responsibility for warning that storm in its own CWA, resulting in two or more warning polygons on one storm that may be issued at nonsequential times (Fig. 3b). Inconsistent messaging also occurs when sections of neighboring counties are excluded at CWA borders (Fig. 3c) or when two WFOs issue warnings of a different type on their respective side of the CWA border (Fig. 3d).

These discontinuities in warning polygons may stem from differences in WFO service needs and office cultural and technological barriers. WFOs focus on balancing tradeoffs between maximizing the probability of detection, minimizing false alarm ratio, and maximizing lead time (Brooks and Correia 2018; Beitlich et al. 2019). These discontinuities are manifested through differences in warning duration, warning on a specific hazard or all hazards accompanying a storm, and product adaptation to end user and community needs (Harrison and Karstens 2017; Klockow-McClain et al. 2020, 2021). The diversity of WFO cultures may come from these needs but also draws from the individual makeup of staff experience,

Corresponding author: Joseph James, jjj27@uakron.edu

DOI: 10.1175/WAF-D-21-0120.1

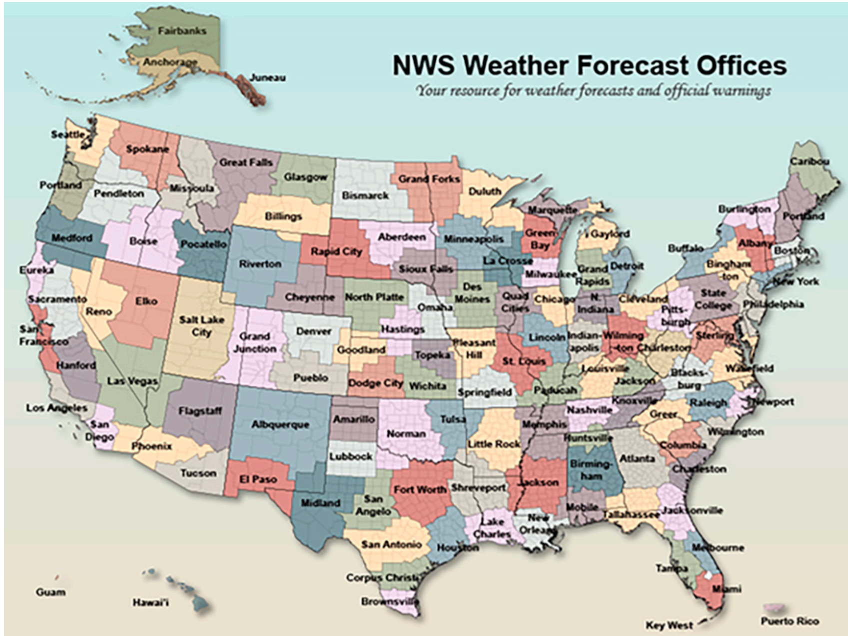

FIG. 1. NWS County Warning Areas for the United States (Source: NWS 2021).

severe-storm expertise, management style, and other factors. Finally, even if WFOs agreed on every warning decision, NWS policy requires warnings to terminate at CWA borders, limiting the possibility of seamless warnings across borders.

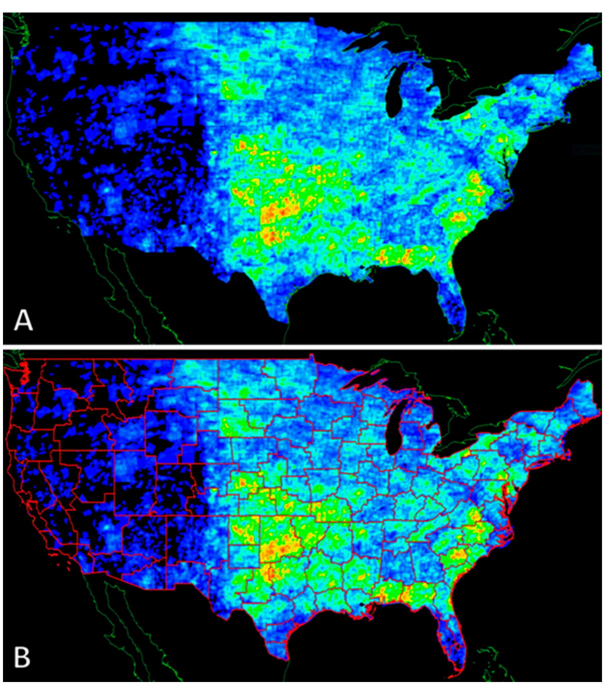

FIG. 2. Severe thunderstorm warning density on a 1-km^2 grid for the period 2015–17. Warmer colors (yellow, orange, red) indicate a higher density of warnings; black indicates no severe thunderstorm warnings for those areas. (a) CWA borders are not shown. (b) CWA borders are shown in red.

To address these shortcomings, Forecasting A Continuum of Environmental Threats (FACETs; Rothfusz et al. 2018) is an experimental framework developing innovative ways to convey rapidly updating weather hazard information from days before to within minutes of an event. This study focuses on the severe weather warning scale (0–2 h, ~100 km^2). Integral to this scale is the creation, management, and communication of gridded, probabilistic threat areas that update at rapid intervals (~1 min), based on the forecast locations of severe weather hazard "objects," defined as 2D areas containing nonzero probability information. The spatial and temporal probability profile reflects the confidence a forecaster places upon the storm object to produce the anticipated hazard (Karstens et al. 2018). This results in frequently updating, probabilistic hazard information (PHI) "plumes" that follow the storm objects continuously in time, with a duration as long as the forecaster has confidence that the storm will exist (up to 2–3 h if the storm is forecast to be long lived) (Fig. 4). The same set of forecasted, 2D storm objects are used to derive warning polygons [i.e., threats-in-motion (TIM); Stumpf and Gerard 2021]. Although these warnings share the same initial 2D object as the PHI plume, the warnings are defined by time and are not tied to probability values. The warning durations are defined to be equal to or less than the plume duration, typically 45 (30) minutes for severe thunderstorm (tornado) warnings, similar to current default NWS warning durations (Brooks and Correia 2018). If the storm is expected to exist beyond the warning duration, the warning polygons can be set to "persist," meaning they will continuously move with the storm at 1-min intervals.

The PHI concept and early prototypes have been under development by the NOAA NSSL and evaluated within the NOAA Hazardous Weather Testbed's (HWT) Experimental

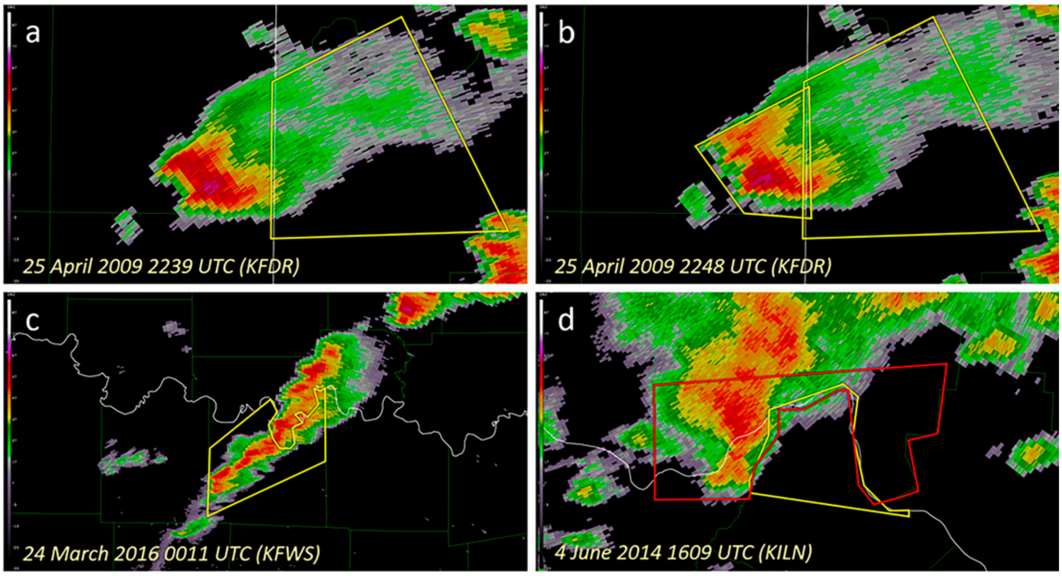

FIG. 3. (a),(b) Images are severe thunderstorm warning polygons (yellow) for a single storm crossing a CWA border. The image in (a) is at an earlier time (2239 UTC). The image in (b) is at a later time (2248 UTC)—the CWA border is between both polygons (thin white line). (c) Image shows a truncated severe thunderstorm warning polygon (yellow) at the edge of a CWA border (thin white line—the Red River). (d) Image depicts a severe storm straddling a CWA border (thin white line —the Ohio River), and an associated tornado warning polygon (red) and a severe thunderstorm warning polygon (yellow), each issued by two respective WFOs. Each image shows the date, time, and WSR-88D radar four-letter identifier in the bottom left.

Warning Program since 2008 (Stumpf et al. 2008; Kuhlman et al. 2008; James et al. 2020; Calhoun et al. 2021). A newer PHI prototype has been undergoing evaluation within the HWT since 2014 (Karstens et al. 2015, 2018). Its concepts and functionality are being integrated into an experimental version of AWIPS II Hazard Services (HS), the NWS's next-generation, object-based, hazard information-generation system that incorporates and will eventually replace WarnGen (Argyle et al. 2017). Named HS-PHI, this software has been tested during simulated WFO warning operations in the HWT since 2016 and is used for this study (Stumpf et al. 2018, 2020; James and Ling 2019).

The continuously updating storm objects and associated plumes and warnings move independently of geopolitical borders, leading to PHI plumes issued by one WFO extending across one or more of its neighboring CWA borders. HS-PHI was designed to reflect this paradigm and thus constructed to aid in potential interoffice collaboration. Collaboration is facilitated through the use of storm object management tools that enable seamless information flows across borders and the ability to view real-time edits from any forecaster. The potential to issue hazard information in this manner, facilitated by new technology, represents a significant departure from the current concept of operations in how adjacent WFOs interact with each other during severe weather. We hypothesize that these software functionalities within HS-PHI may encourage forecaster collaboration and coordination when PHI objects are near adjacent WFO borders. This increased collaboration aligns with the FACETs goal of seamless messaging across CWA borders and could result in more consistent hazardous weather messaging.

Previous HS-PHI HWT evaluations included an exercise to study interoffice collaboration when using PHI and warnings (Klockow-McClain et al. 2020). Two forecasters, representing two adjacent WFOs, staffed two separate workstations in opposite corners of the HWT operations area and used a chat room tool as their sole means for "interoffice" communication as they made decisions on the ownership and handoff of threat objects that were affecting CWA boundaries. Object ownership and handoff were mutually agreed upon by the two forecasters, as the HS-PHI software had not yet included the technology to manage and lock storm objects that move between multiple WFOs.

Forecaster interviews in our earlier work clearly demonstrated that interoffice collaboration could present itself as a major barrier to a successful implementation of a FACETs-based warning program (Klockow-McClain et al. 2020). This paper aims to fulfill the need to conduct a more formal study of interhuman collaboration within the context of the warning software. One significant challenge of this ongoing research is investigating a hazard-information-alerting system with continuous storm objects that are seamlessly transferred between WFOs via collaboration tools.

While research on human-to-human collaboration within automated system functions is relatively new in the area of weather forecasting, handoff and collaboration is well studied in the field of aerospace regarding air traffic controllers (ATCs). ATCs regularly transfer the responsibility of aircraft

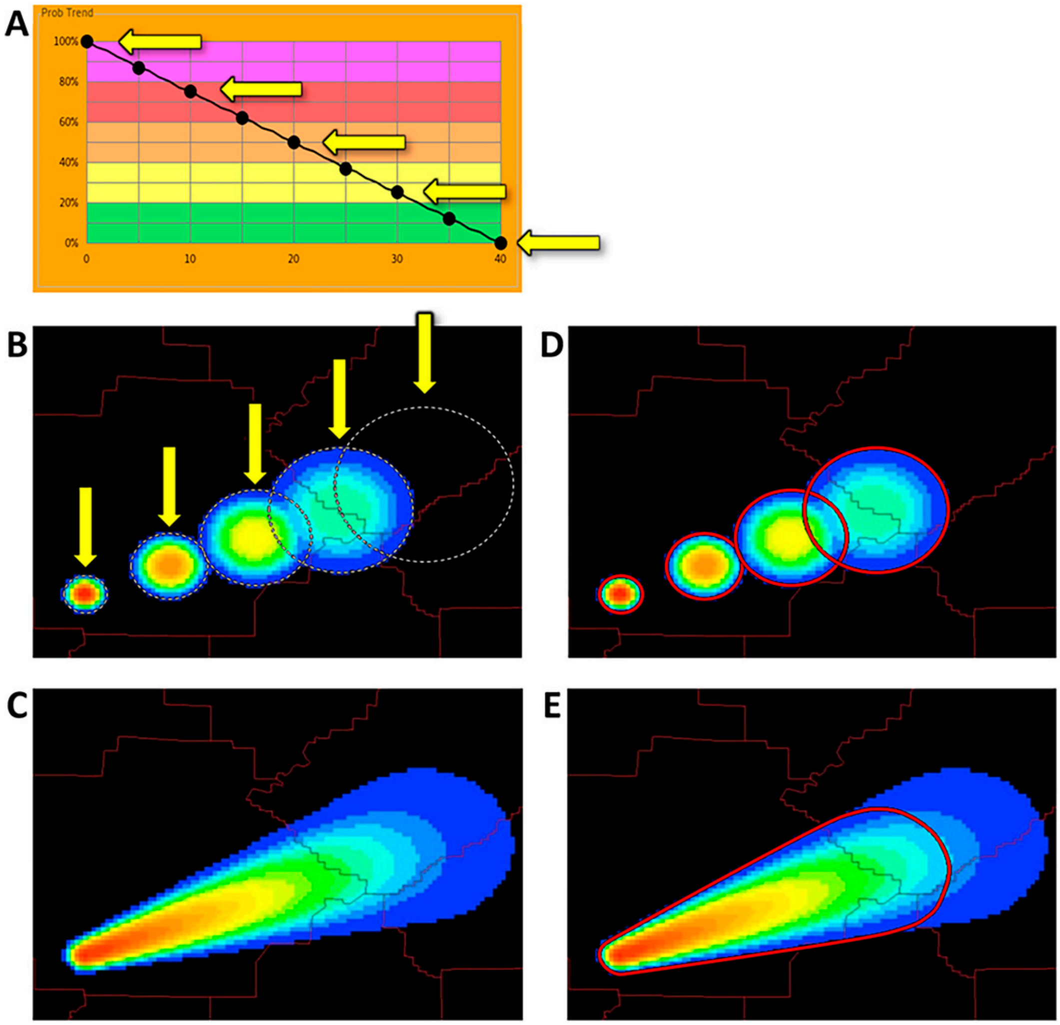

FIG. 4. An illustration of the process to create PHI plumes and warnings from 2D storm objects that are projected to 1-min forecast objects. Only the 10-min forecast objects are shown (time = 0, 10, 20, 30, 40 min). (a) The forecaster-defined probability trend, linearly decreasing from 100% to 0%, for a plume duration of 40 min, with arrows indicating values every 10 min. (b) The forecast probability objects with values decreasing from the maximum probability trend values at the center of the object to 0% at the edge of the object (dashed outline), with arrows indicating values every 10 min (at 40 min, probability = 0%). (c) The maximum probability value of all 1-min forecast objects defines the PHI plume. (d) The outline of forecast objects for a warning duration of 30 min (defined to be shorter than the plume duration), and (e) the cascaded union of all 1-min forecast objects creates the warning polygon.

from one ATC to another through a handoff process that can be completed using built-in system capabilities (FAA 2015). Landry et al. (2010) found that in a collaborative environment across multiple ATCs, common data display improved alignment on decisions and facilitated collaboration. A similar concept has been implemented into the collaboration aspect of the current experiment. When evaluating workload in a complex ATC environment, latencies in handoff initiation and acceptance can cause disruption in the workflow (Rantanen 2004). Similarly, during handoffs between neighboring WFOs in severe weather operations, latencies in handoff and acceptance may cause workflow disruption.

In potential future severe weather operations, PHI objects could likewise be transferred directly between neighboring WFOs. Drawing from past research on ATC collaboration (Sharples et al. 2007), considerations need to be made regarding mental workload, handoff strategies, and delays in the PHI handoff process. In this study, we investigate how forecasters collaborate across CWA boundaries, given that they potentially have different warning philosophies and workload, and how the performance of these new tools impact their collaboration and workload. Note that warning verification was not investigated in this study. Specifically, our research questions to be addressed include:

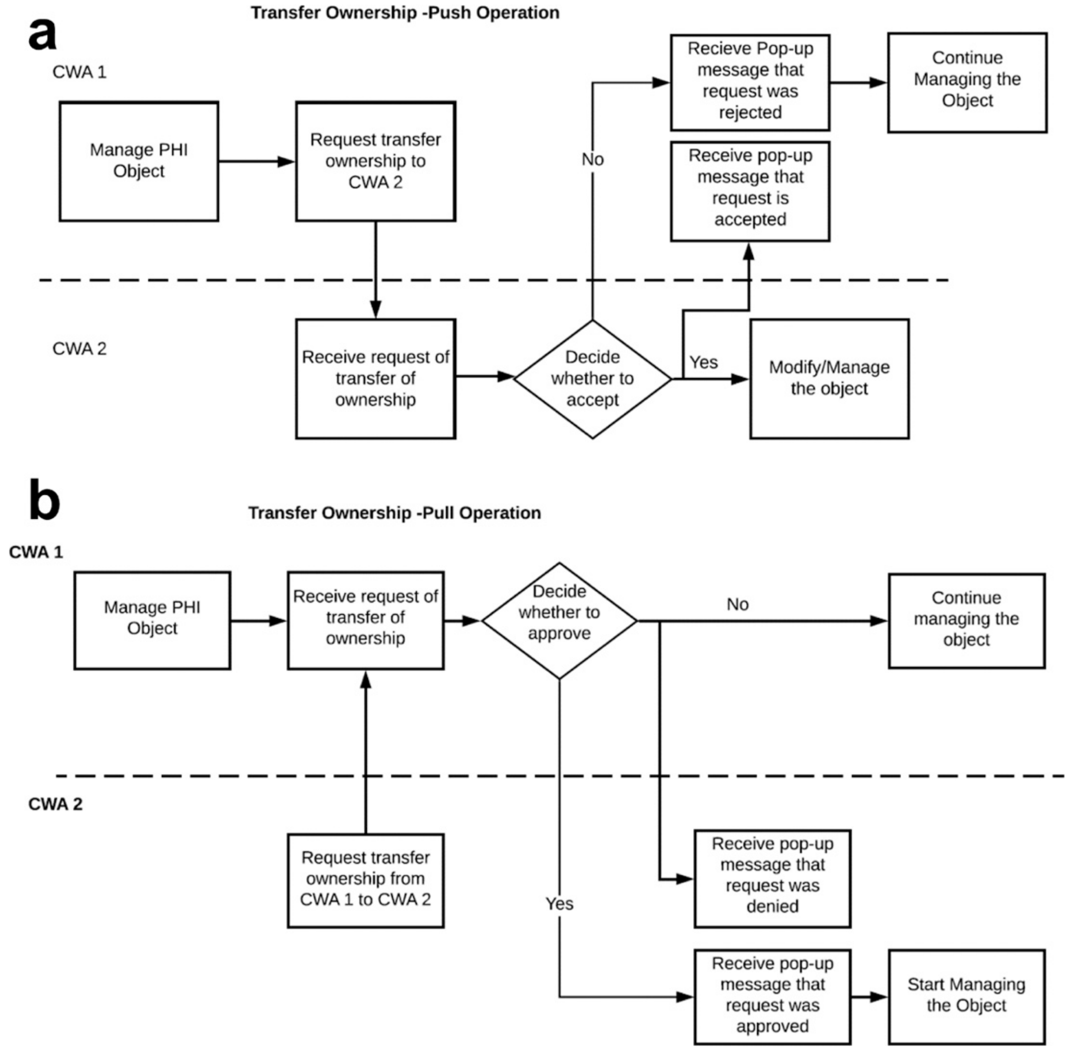

FIG. 5. Transfer of ownership push operation. Forecasters followed this decision flowchart when requesting an adjacent CWA take over ownership of one of their objects. (a) Flowchart shows a push operation, when a forecaster wants to give control of an object and (b) a pull operation is shown, when a forecaster is requesting to take over an object.

- How do forecasters collaborate?
- How do they collaborate when different philosophies exist among forecasters?
- How do they feel about the performance of the collaboration tools?
- How hard do they need to work to collaborate and issue warnings?

2. Methodology

a. HS-PHI collaboration software functionality

The 2020 experiment tested new tools and functionality designed within HS-PHI to facilitate ownership and handoff of storm objects. When a storm object was created, the forecaster who produced the object maintained ownership and editing control until the object's expiration or transfer of ownership to another forecaster in an adjacent CWA. New to HS-PHI for this experiment, transfer of ownership, or "handoff" (of an object), could be accomplished in one of two ways: 1) The forecaster owning the object could suggest a transfer of the object to another forecaster through a "push" operation, or 2) A forecaster could ask for a transfer of an owned object not belonging to the forecaster through a "pull" operation. In each case, the forecaster receiving the transfer request must either "accept" or "deny" the transfer request. Forecasters had the ability to communicate with adjacent WFOs via the objects' warning decision discussions, chat, and phone calls to discuss object transfers.

Figure 5a shows a push ownership transfer operation, where CWA1 owns a storm object and requests to transfer ownership to CWA2. The CWA1 forecaster selects the object and workstation to send the transfer request, as shown in Fig. 6a. After the request is sent, an accept or decline message appears on the receiving forecaster's workstation screen, as shown in Fig. 6b.

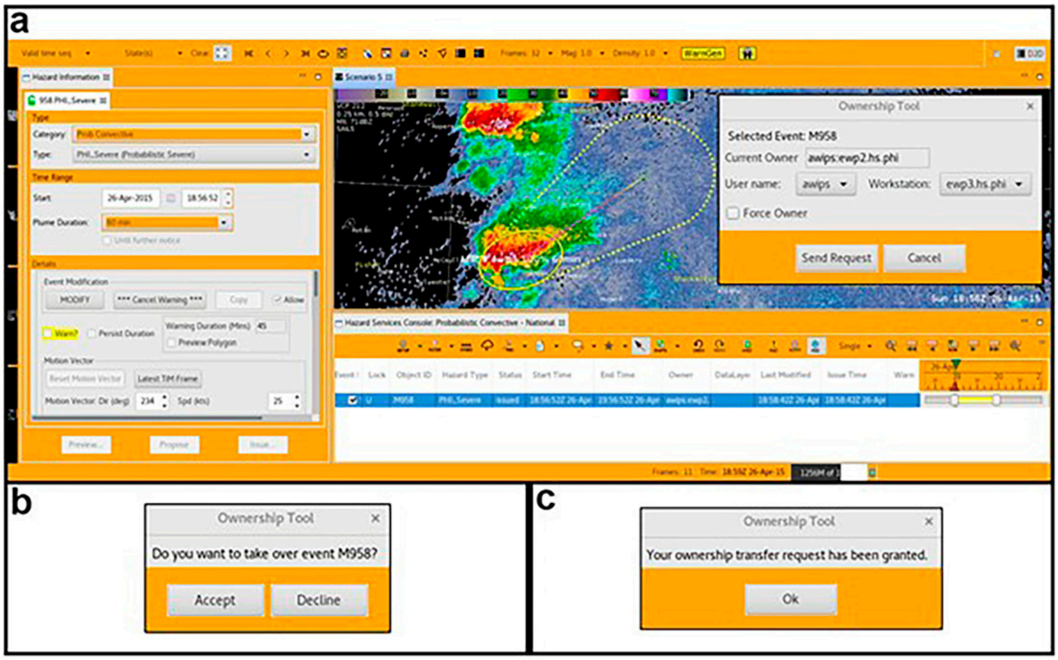

FIG. 6. HS-PHI images during the transfer-ownership push operation that is illustrated in Fig. 5. (a) The HS-PHI AWIPS display configuration showing a single storm object (yellow ellipse), proposed warning plume (yellow dashed line), and the ownership tool pop-up window that allows the WFO in CWA1 to send a transfer request to the WFO in CWA2. (b) After the WFO in CWA1 sends the request, this pop-up window is seen on the workstation of the WFO in CWA2 that alerts them of the receipt of a request of transfer of ownership. (c) After the WFO in CWA2 accepts the transfer request, this pop-up window that the request was accepted is seen on the workstation of the WFO in CWA1.

Once the receiving forecaster accepts the transfer request, a confirmation appears in CWA1 and object ownership and editing privileges switch to the receiving forecaster. The pull ownership transfer operation is similar to the push operation, except the initiating forecaster is requesting to take ownership of the storm object from another forecaster, illustrated in Fig. 5b.

The domain permission tool (Fig. 7) adds flexibility for forecasters to create storm objects that straddle a border, potentially outside the bounds of their CWA. If a forecaster wants to create an object that extends beyond the border of their CWA into an adjacent CWA, or in the case that a storm is encroaching on their CWA border but not quite within their CWA, the forecaster has to request permission first from the neighboring CWA to create an object. This domain permission request, only used near CWA borders, is automatically initiated when a forecaster attempts to create an object that overlaps a CWA boundary. A message appears on the screen of a forecaster in the neighboring CWA, similar to the ownership transfer operations, with the ability to accept or decline domain permission.

In addition, HS-PHI allows forecasters to view real-time edits to PHI objects, plumes, and warnings from any of their neighboring WFOs. HS-PHI also allows forecasters to add textual warning decision discussions to each storm object, providing neighboring WFOs with a "written history" of forecaster's thoughts for those storms. All of these features provide collaborative capabilities that currently do not exist in WarnGen.

b. *Participants*

For the 2020 experiment, a nationwide solicitation for applications from NWS forecaster pairs was disseminated, and preference was given to pairs from the same WFO. With two pairs per week and two experiment weeks, pairs from the same WFO were kept together during the experiment. Overall, a total of eight forecasters participated in the experiment, with warning-operations experience ranging from 2 to 16 years.

c. *Experimental design*

The HWT experiment was designed to replicate the sociocultural conditions that may lead to challenges in interoffice collaboration in the present deterministic system while also exploring the ways those influences would affect judgments and collaboration in a probabilistic system. The overall experiment was designed based on information gleaned from WFO interviews regarding warning policies, relationships with core partners, and other relational and policy considerations (Klockow-McClain et al. 2021).

d. *Daily activities*

On the first day of the weeklong experiment, participants were presented with introductory materials and an experiment

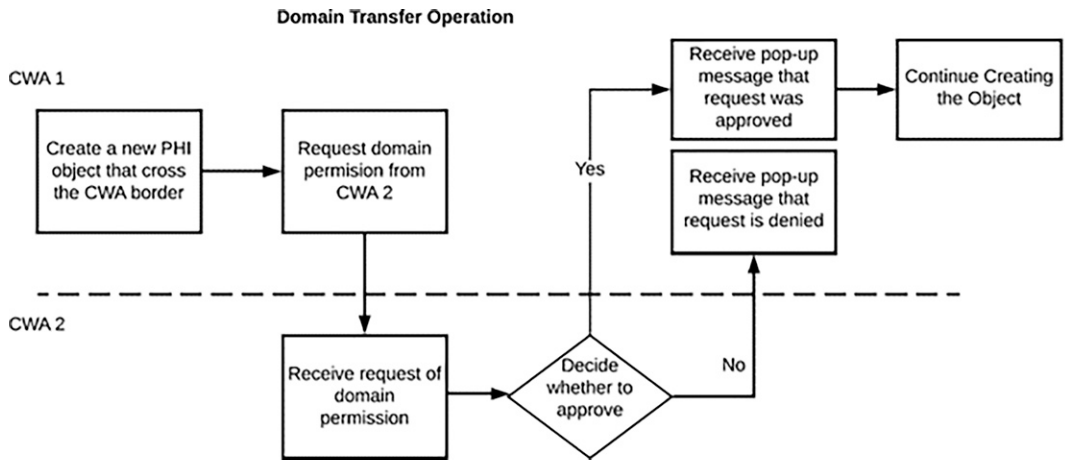

FIG. 7. Domain transfer request decision flowchart, when forecasters want to create an object that partially overlaps a CWA boundary, forecasters must ask for permission from adjacent CWA forecasters for approval.

overview. Following the briefings, forecasters received hands-on training via job sheets to build familiarity with the HS-PHI software and tools, with assistance from project scientists and facilitators. The following three days, forecasters completed two scenarios each day using HS-PHI for severe weather threats. The seventh scenario ran the full day with a session in the morning and a continued session in the afternoon.

Forecasters were split into two rooms to represent two WFOs with two forecasters per WFO. Each room had two dual-display AWIPS II workstations and situational awareness monitors, similar to previous Hazard Services experiments conducted in the HWT (James et al. 2020). In addition to the workstation setup, forecasters had access to a standard conference phone system and a tablet/keyboard combination with access to a proprietary business communication software to communicate with the other testing laboratory (“adjacent” WFO).

Forecasters worked together on severe weather scenarios, communicating and collaborating as they saw fit on storms crossing CWA boundaries. Workload and collaboration surveys were completed after each scenario. Guided intra- and interoffice (referring to the fictitious WFOs within this experiment) discussions took place following each scenario and survey. The guided sessions covered topics of interest that resulted from forecaster decisions/actions or situations that occurred during the experiment scenarios. Following four days of experiment scenarios, project scientists and facilitators conducted a comprehensive, guided discussion with the forecasters for feedback regarding the HS-PHI software and collaboration tools, experiment design, and the future of warning operations in this potential new framework.

e. Scenario information

Experiment scenarios were carefully selected to test collaboration strategies along CWA borders. These scenarios utilized current and modified CWA borders, and included storms moving from one CWA to the other, storms straddling CWAs, and storms impacting multiple CWAs simultaneously due to having large spatial extents (e.g., squall lines). A variety of storm modes and variations of storm interactions along CWA borders were present in the scenarios. Five different scenarios were chosen for the experiment (Table 1).

To replicate NWS operations further, the scenarios included office directives (from researchers acting as WFO management) to simulate differing, and conflicting, warning philosophies that could result in warning discontinuities similar to those in Fig. 3. For example, a simulated WFO was directed to give their severe thunderstorm warnings abnormally long durations, causing the warnings to extend into the neighboring CWA. Fictional Impact-based Decision Support Services (IDSS) events (Uccellini and Ten Hoeve 2019), such as emergency managers supporting a sailboat race along a river, were strategically placed near CWA borders during the scenarios, with interjections during warning operations for the forecasters to answer site-specific questions. The complexity of the scenarios increased throughout the week as more IDSS events and directives were added, culminating in a daylong simulation. Additionally, to test neighboring CWA awareness, forecaster pairs rotated roles as either the upstream or downstream office during the experiment.

f. Human factors analysis

1) NASA-TLX MENTAL WORKLOAD QUESTIONNAIRE

The NASA-TLX (Hart and Staveland 1988; Hart 2006; James et al. 2020; James 2021) workload index is a questionnaire-based workload rating tool. The tool encompasses six aspects of workload: mental demand, physical demand, temporal demand, performance, effort, and frustration. Modifications were made to the questionnaire in the form of a follow-up question to ascertain the reasoning for each rating. The raw scores of the mental workload range from 1 to 100, with 1 for extremely low workload and 100 for extremely high workload. Each pair was cross-compared and forecasters selected which pair was more “important.” This comparison was averaged to find the “importance” factor of each mental workload aspect. The scores were averaged for all forecasters overall, forecaster

TABLE 1. A summary table of scenario attributes, including which WFOs the forecasters were operating in, the date/time of the case that was used, the meteorological characteristics of the event and warning type used, number of fictional IDSS events per WFO, and number of mid-scenario directives per WFO.

Localized NWS WFO	Time/date	Meteorological description	No. of fictional IDSS events per WFO	No. of directives per WFO
San Angelo and Ft. Worth, TX	1840–2040 UTC 26 Apr 2015	Isolated supercells crossing vertical and straddling horizontal CWA borders (SVR and TOR)	SJT: 1 FWD: 1	SJT: 0 FWD: 0
Duluth and Minneapolis, MN	1930–2200 UTC 19 Jul 2019	Line of supercells straddling CWA border horizontally, marginally tornadic (SVR and TOR)	DLH: 1 MPX: 1	DLH: 0 MPX: 0
Birmingham, AL, and Atlanta, GA	1900–2100 UTC 3 Mar 2019	Supercells and a QLCS, long-tracked tornado crossing vertical CWA border (SVR and TOR)	BMX: 0 FFC: 1	BMX: 1 FFC: 1
Modified Albany NY	1750–2020 UTC 5 May 2018	Squall line and supercells along and crossing vertical CWA border (SVR and TOR)	West: 4 East: 3	West: 1 East: 1
Norman and Tulsa, OK	0045–0545 UTC 20 Oct 2019	Supercell and multicell mergers along and crossing vertical CWA border (SVR and TOR)	OUN: 2 TSA: 4	OUN: 1 TSA: 1

pairs, and scenarios for each of the six aspects of workload (James and Ling 2019; James et al. 2020; James 2021).

2) COLLABORATION SURVEY

Collaboration surveys, administered at the end of each scenario as well as at the conclusion of the weeklong experiment, were developed to gauge forecasters' perceptions of the collaboration tools in HS-PHI. These surveys utilize a seven-point Likert scale (Likert 1932) (Table 3). The goal of each end-of-scenario survey was to understand forecasters' perceptions of the effectiveness of collaboration for each scenario, while the goal of the end-of-week survey was to compare the effectiveness of collaboration within the HS-PHI experiment to collaboration strategies utilizing the WarnGen system.

3) THEMATIC ANALYSIS

Thematic analysis is a method often used in psychology for developing patterns and themes within sets of qualitative data (Braun and Clarke 2006). The end-of-week discussion, collaboration survey results, and survey comments were analyzed using thematic analysis. A theme was identified when several forecasters discussed a similar topic; themes usually showed a common usage or process in the way the forecasters used the system or completed tasks. Challenges were identified when forecasters mentioned and noted common issues that prevented them from effectively completing tasks. Themes and challenges were also developed using researcher observations, survey results, and group discussion. Themes identified patterns in forecaster actions and strategies, while challenges identified common issues that needed improvement or solutions, whether for operations or for software. These data were coded and analyzed using Nvivo 10 qualitative data analysis software (Jackson and Bazeley 2019).

3. Results and discussion

a. Thematic analysis results

We observed that when forecasters collaborated over storm objects, plumes, and warnings, forecasters generally used a three-stage strategy: communication, handoff, and post-handoff. Forecasters usually communicated via chat or phone before initiating a handoff, unless they were directed by researchers not to communicate. Then forecasters implemented the handoff with functions provided by the HS-PHI software. The receiving forecaster could then choose to modify the storm object after the handoff was completed. Themes and challenges related to collaboration were derived for each stage of storm object handoff, shown in Table 2. (The number of times a theme or challenge was mentioned is noted in the table.)

1) COMMUNICATION STAGE

Four themes and three challenges emerged to describe the forecaster's tasks during the communication stage. Themes in the communication stage are described as follows.

- Increased frequency of communication

Forecasters mentioned an increase in the frequency of communication as compared to their current WFO operations. This result is likely due to the nature of the experimental setup, as forecasters were encouraged to communicate with the other experimental WFO as they saw fit. Additionally, the scenarios focused on events with a significant amount of storm interactions with CWA borders.

- Allowed for more concise communication by providing storm object, storm motion, and warning-decision discussions and facilitated forecasters from both WFOs to communicate effectively and consistently

While using the HS-PHI software, forecasters mentioned that they were able to pinpoint and discuss specific storm objects easily for ownership transfers, probability-trend philosophies, and other topics over chat and phone. In this regard, it was easier for forecasters to understand the neighboring forecasters' thought processes during warning operations with HS-PHI. Forecasters often used storm object attributes when collaborating. Furthermore, in some instances, forecasters used their neighboring WFO's storm object information to address fictional IDSS requests posed by researchers.

TABLE 2. Themes and challenges of collaboration for each of the three stages of storm object handoff. (Note: "F" denotes the number of times forecasters referenced this in a survey, "D" denotes the number of times forecasters mentioned this in a post-scenario discussion, and "O" denotes the number of times researchers observed this.)

Stages	Theme	Challenge
Communication	1. Increased frequency of communication (O3, F10, D1)	1. Phone and chat collaboration required extra workload, but forecasters acknowledged this was necessary (F2, D1)
	2. Reduced unnecessary communication by providing storm object and storm motion, allowed forecasters from both WFOs to communicate effectively and be on the same page (F3)	2. Being aware of objects about to come into your CWA is sometimes challenging (O1, F1)
	3. Visualization of neighbor CWA's thinking and a quick review of data facilitated chat and phone collaboration (O1, F13, D1)	3. Collaboration is challenging when phone or chat not available (F1, D1)
	4. Chat or phone collaboration is necessary before an object transfer or ownership request is initiated in the tool (F7)	
Handoff	1. The HS-PHI tools offered a quick and easy way to transfer ownership (F4, D2)	1. High workload results in forecasters delaying transfer (O3, F1, D2)
		2. Forecasters need a guideline on when to do a handoff (when an object is touching the border? Half over the border?) (O4, D1)
Post-handoff	1. Receiving forecaster often updates the object immediately after the transfer (O3, F1, D1)	

- Visualization of neighboring CWA's thinking and a quick review of data facilitated chat and phone collaboration

Forecasters frequently mentioned that the visualization of storm objects helped them review data quickly, which was particularly useful when communicating with the neighboring WFO. This was especially true when the neighboring WFOs had differing warning philosophies. The HS-PHI software allowed forecasters to see the warning philosophies of their neighboring WFO (via the warning-decision discussions appended to each object) and use that information as a basis for initiating a phone call or chat to collaborate on storm objects. This capability is not available in the current WarnGen framework. A forecaster during the second week of the experiment stated:

> "The software allowed me to easily identify philosophy differences, which then gave me the option to communicate with the neighboring CWA if desired. In the current WarnGen paradigm, I get a sense of my neighboring WFO's warning decision process by observing their SPS vs warning vs nothing decisions on the radar. With HS-PHI, I get the same thing by means of plumes vs warnings vs nothing, but also get extra information by examining their probabilities, discussions, and TIM decisions. Better!"

- Chat or phone collaboration is necessary before an object transfer or ownership request is initiated in the tool

Throughout the experiment, discussion via chat or phone typically occurred before storm objects were transferred. When a forecaster received a transfer request without previous communication, confusion and object-transfer delays resulted. In an instance where researchers directed the downstream forecasters to be unresponsive, an object transfer was delayed by at least 5 min.

When forecasters communicated before handoff, they also encountered various challenges. Three challenges were mentioned frequently by forecasters and researchers.

- Phone and chat collaboration required extra workload, but forecasters acknowledged this was necessary

Forecasters realized that collaboration over the phone and chat resulted in increased workload, but they understood this step in the collaboration process was crucial to achieve the continuous flow of seamless information.

- Situational awareness of inbound-CWA PHI objects can be challenging

Forecasters, especially in the downstream WFO, found it challenging (increased mental workload) to manage storm objects within their own CWA while also being cognizant of objects that were approaching their CWA from the upstream WFO.

- Collaboration is challenging when phone or chat not available

Since a significant amount of storm-object transfers were initiated over chat or a phone conversation, when one WFO was unable to contact the other WFO or did not receive a response, the transfer or exchange of information on an object was hindered. Some of this unresponsiveness was the result of experimental directives by the project scientists as a method to mimic a busy or uncooperative WFO.

2) HANDOFF STAGE

During the handoff stage, one theme and two challenges emerged to describe the forecasters' tasks. The theme in the handoff stage is described as follows.

TABLE 3. The 2020 end-of-scenario collaboration survey results averaged across all CWA's and scenarios (1 being strongly disagree and 7 being strongly agree). SD refers to standard deviation.

End-of-scenario collaboration question	Mean/SD
1) With the HS-PHI software, I felt aware of the neighboring WFO's warning decision process.	4.9/1.3
2) I felt comfortable when a PHI plume existed in my CWA and the neighboring CWA at the same time.	5.9/0.9
3) I felt comfortable when my neighboring office had control of a PHI plume that was also in my CWA.	5.4/1.3
4) I felt comfortable when my neighboring office made a decision that modified the PHI plume within my CWA.	5.3/1.3
5) I felt comfortable when a warning existed in my CWA and the neighboring CWA at the same time.	5.8/1.1
6) I felt comfortable when my neighboring office had control of a warning that was also in my CWA.	5.2/1.2
7) I felt comfortable when my neighboring office made a decision that modified the warning within my CWA.	5.0/1.3
8) Tracking hazards using 2D objects helped me to collaborate with my neighboring WFOs.	5.4/1.4
9) The ownership transfer tool and the domain permission tool helped me to collaborate with my neighboring WFOs.	5.3/1.8
10) The HS-PHI software fostered effective collaboration between me and my neighboring WFO.	5.7/1.4
11) The HS-PHI software fostered effective collaboration between me and my neighboring WFO when our warning and/or PHI philosophies differed.	4.6/1.5
12) The HS-PHI software fostered efficient collaboration on PHI plumes and warnings.	5.3/1.5
13) I easily developed strategies for collaboration when issuing PHI plumes and warnings using the HS-PHI software.	5.9/1.3
Total average	5.4

- HS-PHI tools offered a quick and easy way to transfer ownership

The HS-PHI collaboration tools provide a means for forecasters to collaborate and create seamless PHI and warnings across CWA borders. Once an agreement on an object transfer occurred, forecasters found the tools enabled them to complete the transfer quickly and easily. A forecaster stated, "When you do a handoff, the transition is longer with a larger object, whereas the transfer is quicker with smaller objects."

The handoff process presented two notable challenges, as described below

- High workload results in forecasters delaying transfer

Delayed object transfers due to high workload occurred when downstream forecasters, who sometimes already owned many objects, received a transfer request for a border object from the upstream WFO. Downstream forecasters either would wait to accept an object or would request the upstream WFO to manage an object, sometimes well into their own CWA, until workload diminished.

- Forecasters need guidance for handoffs in proximity to CWA borders (when an object is touching the border? Half over the border?)

The process of transferring objects from one WFO to another is subjective. There is no guidance for when objects should or could be transferred to another WFO in this new framework. Forecasters discussed the challenge of when it is appropriate or required to transfer an object. Several variables were mentioned as factors for consideration, including workload, object size and motion, and societal factors (such as population centers or IDSS activities).

3) POST-HANDOFF STAGE

One theme was mentioned frequently by both forecasters and researchers during the post-handoff stage.

- The receiving forecaster often updated the object immediately after the transfer.

After an object's ownership transfer, the receiving forecaster often updated, altered, or ended the storm object. These changes by the downstream WFO occurred during roughly 50%–75% of object transfers, workload permitting. If the receiving forecaster wanted to end the transferred object, they did not do so immediately. Instead, the receiving forecaster maintained the object for a short period of time (roughly 10–15 min) before ending the object. Transferred objects that are quickly and substantially changed by another forecaster after a handoff can be problematic. For example, immediately changing a long-duration warning to a shorter-duration warning, or vice versa, could evoke confusion with end users and publics who are suddenly removed from or added to a warning. One forecaster stated, "There were times when we didn't quite agree, so when the transfer of an object was done, a couple of changes were made, maybe not right away, but quick enough. I can see this being an issue at times in an operational environment." Forecasters further discussed conflicting messaging to the public resulting from quick changes by forecasters who disagreed.

b. *Collaboration survey results*

One survey was administered at the end of each scenario to evaluate collaboration effectiveness. Another survey was administered at the end of the week to compare collaboration effectiveness with the current WarnGen warning method. Overall, the average rating for the end-of-scenario collaboration survey was 5.4 out of 7 (median, 6; range, 6) (slightly agree) (Table 3) and the average rating for the end-of-week collaboration survey was 5.0 out of 7 (median, 5; range, 6) (slightly agree) (Table 4), showing a generally positive attitude toward the collaboration effectiveness provided by the HS-PHI software. One possible reason that this rating is slightly lower than those of the end-of-scenario ratings could

TABLE 4. The 2020 End-of-week collaboration survey results. SD refers to standard deviation.

End-of-scenario collaboration question	Mean/SD
1) With the HS-PHI software, I felt more aware of the neighboring WFO's warning decision process.	4.8/1.0
2) Tracking hazards using 2D objects improved my ability to collaborate with my neighboring WFOs.	5.0/1.0
3) The ownership transfer tool and the domain permission tool improved my ability to collaborate with my neighboring WFOs.	5.3/1.8
4) The HS-PHI software fostered more effective collaboration between me and my neighboring WFO.	5.1/1.8
5) The HS-PHI software fostered more effective collaboration between me and my neighboring WFO when our warning and/or PHI philosophies differed.	4.3/1.5
6) The HS-PHI software fostered more efficient collaboration warnings.	5.1/1.8
7) I can develop better strategies for collaboration when issuing PHI plumes and warnings using the HS-PHI software.	5.3/1.2
Overall average	5.0

be because the last scenario featured a long-duration event with many storms.

For the end-of-scenario collaboration survey, forecasters found the highest agreement with statements (average score denoted in parentheses), "I felt comfortable when a warning existed in my CWA and the neighboring CWA at the same time. (5.8)" and "I easily developed strategies for collaboration when issuing PHI plumes and warnings using the HS-PHI software. (5.9)." The first statement is significant because, in the current warning paradigm, warnings are not allowed to cross borders (warnings are clipped/truncated at CWA borders). The second result shows forecasters were able to develop collaboration strategies quickly, though forecasters had very limited experience with the new software and PHI paradigm.

Forecaster agreement was lowest for the statement, "The HS-PHI software fostered effective collaboration between me and my neighboring WFO when our warning and/or PHI philosophies differed. (4.6)." While this is still on the side of agreement, this result received a low rating mainly because of the condition of "when our warning and/or PHI philosophies differed" in this statement (researchers imposed differing warning philosophy directives throughout the experiment). In contrast, a similarly worded question without this condition received an average rating of 5.7. These results show that in situations where neighboring offices have different philosophies, forecasters have difficulties collaborating across WFOs. When comparing average collaboration agreement levels between probabilistic and traditional warning paradigms, they were very similar; however, collaboration agreement regarding warnings with probabilistic information was slightly higher (5.5 compared to 5.3, respectively). The collaboration agreement was very similar for both upstream and downstream WFOs.

For the end-of-week collaboration survey, the statements with the highest agreement were, "The ownership transfer tool and the domain permission tool improved my ability to collaborate with my neighboring WFOs. (5.3)" and, "I can develop better strategies for collaboration when issuing PHI plumes and warnings using the HS-PHI software. (5.3)." A forecaster during the second week of the experiment stated, "The domain permissions tool was helpful in knowing what the other WFO was trying to do in our CWA and knowing that collaboration may be needed." These results, illustrating improved collaboration in this probabilistic framework compared to the current framework, are analogous to the results of the individual post-scenario surveys. One collaborative software feature that was favorable among forecasters was the ability to watch a forecaster in their neighboring WFO create PHI and warning objects in real-time, a capability that does not currently exist in WarnGen.

Forecasters were least agreeable with the statement, "The HS-PHI software fostered more effective collaboration between me and my neighboring WFO when our warning and/or PHI philosophies differed. (4.3)," for the end-of-week survey. As before, this is mainly because of the question phrasing. A similarly worded question without the aforementioned condition received an average rating of 5.1. One forecaster commented that the HS-PHI software fostered collaboration by allowing him to view his neighbor's thought process. These results show that although forecasters felt that interoffice collaboration was improved in a probabilistic framework compared to current operations, in specific situations, the software alone was not enough to support collaboration when neighboring forecasters disagreed on probabilities or a warning decision.

Combined, the collaboration survey results show overall favorable attitudes toward the software and probabilistic paradigm, and support the hypothesis that the software additions encouraged forecasters to be more collaborative with their neighbors. Given forecasters' comfort with warnings existing outside the CWA of origin, their ability to quickly develop and easily collaboration strategies using the HS-PHI software is demonstrated. The forecasters' higher confidence utilizing a probabilistic framework compared to the current WarnGen framework is likewise shown. Altogether, a collaborative format within the FACETs paradigm for future NWS operations is feasible.

c. *Collaborative warnings versus NWS warning examples*

The following examples from the collaboration experiment illustrate how seamless warnings can be generated across CWA borders, versus multiple disparate warnings on either side of the borders. Warning verification is not addressed for this discussion. The warnings generated with HS-PHI were CWA-agnostic and a product of two WFOs collaborating on a unified message.

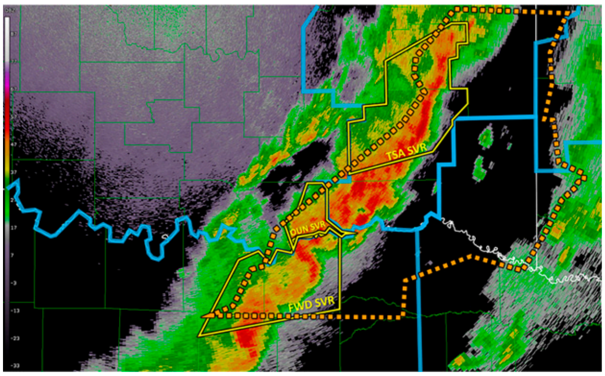

FIG. 8. A screenshot from the 2020 HS-PHI tool showing the experimental PHI-associated severe thunderstorm warning (orange dash outlined area) compared to the operational NWS severe thunderstorm warnings (solid yellow outlined areas) issued by three different WFOs: Tulsa (TSA), Norman (OUN), and Fort Worth/Dallas (FWD). The blue lines are the CWA borders.

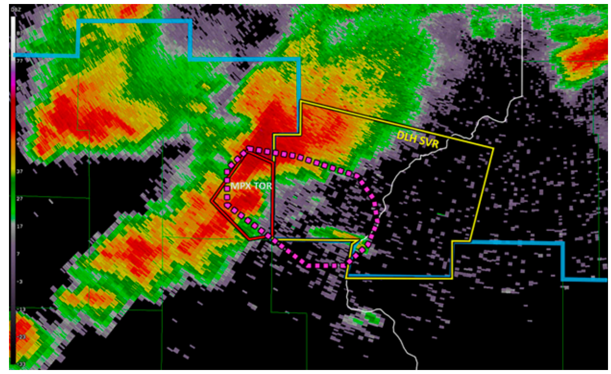

FIG. 9. A screenshot from the 2020 HS-PHI tool showing the experimental PHI-associated tornado warning (pink dash outlined area) compared to the operational NWS tornado warning (TOR; red solid outline area) and severe thunderstorm warning (SVR; yellow solid outline area) issued by two different WFOs: Minneapolis (MPX) and Duluth (DLH). The blue line is the CWA border.

Figure 8 shows an HS-PHI severe thunderstorm warning compared to the three original severe thunderstorm warnings issued by three different WFOs: Tulsa (TSA), Norman (OUN), and Fort Worth/Dallas (FWD). The original warnings are geographically terminated at the respective CWA borders. In contrast, the PHI-based warning covers the entire threat area with one warning and provides seamless and consistent warning information across five CWAs. For this example, we assume that the threat is the same across the entire line—if the threat varied across the line, there could be multiple warning polygons for each segment of the line representing different intensities of threat. Regardless, these multiple warnings would remain seamless across the CWA borders and represent a consistent message from every WFO.

Figure 9 shows a tornado warning created with HS-PHI. The original tornado warning from the NWS Minneapolis WFO (MPX) and the original severe thunderstorm warning (which includes a "Tornado Possible" tag) from NWS Duluth WFO (DLH) are also shown in the image. The NWS warnings terminate geographically at the CWA border, leading to potentially conflicting messaging for those on either side of the border. The HS-PHI tornado warning, however, covers the entirety of the threat area across two WFOs as a unified and consistent message. While the storm may have necessitated a downgrade from a tornado warning to a severe thunderstorm warning, this downgrade would have been a collaborative decision by both WFOs, and likely would not happen exactly at the CWA border.

d. Mental workload

While the experiment attempted to replicate NWS WFO operations as closely as possible, forecasters regularly reported that workload was an issue during the experiment, especially given that the two forecasters per experimental WFO were tasked with handling all warning operations, IDSS events, and collaboration. Although forecasters individually handled multiple tasks, the whole of these responsibilities is typically shared among a larger group of forecasters within WFOs.

The average mental workload for this experiment across all scenarios was 58.8 out of 100, an increase from 54.4 out of 100 reported in the 2019 HS-PHI experiment (James 2021) (Fig. 10). Of the workload subdimensions, mental demand, temporal demand, and effort had the highest ratings among the rest of the workload scores. The latter results show that forecasters experienced a significant workload related to decision-making and other cognitive demands in a fast-paced environment. Such an environment requires tremendous mental effort to complete a variety of experiment-driven tasks that may be outside the scope of their current working environment. Some forecasters even verbalized that they were physically tired after a particularly busy

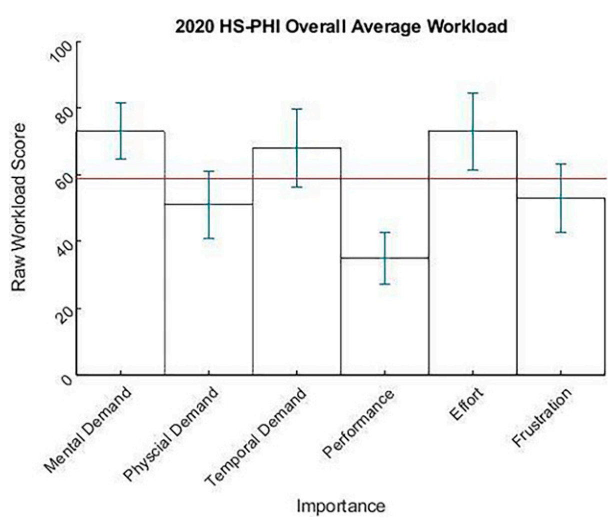

FIG. 10. Average NASA-TLX workload for each subdimension across both CWAs and all scenarios. The red line represents the average mental workload of all five cases. The importance factor (x axis) is the relative importance of each mental workload subdimension.

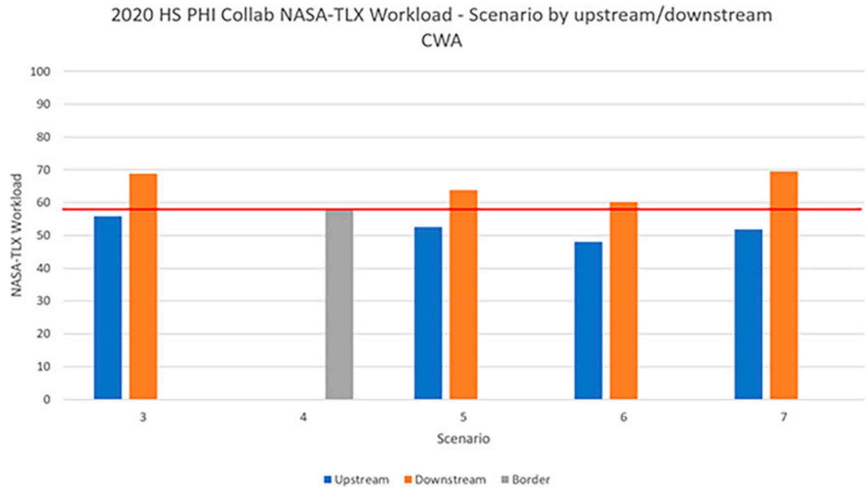

FIG. 11. Average NASA-TLX workload comparing the workload of forecasters in downstream CWAs compared to upstream CWAs per scenario (scenario 4 did not have clear upstream/downstream CWA; a gray bar is used to represent average mental workload). The red line represents the average mental workload of all five cases.

scenario. This was especially true after scenario 7, the full-day event that featured many storms, possibly accounting for the uptick in workload for that scenario.

The downstream CWA consistently reported a higher mental workload (10 points higher than the upstream CWA) across all but one of the scenarios (see Fig. 11; scenario 4 was a border case without a clear upstream/downstream CWA distinction). The upstream CWAs reported an average mental workload of 52, while the downstream CWAs reported an average mental workload of 65.5 out of 100. Downstream forecasters stated they experienced increased mental workload because they were not only responsible for managing PHI plumes and warnings within their own CWA, but they also had to maintain situational awareness of storm objects approaching from the upstream CWA. Upstream forecasters stated they often disregarded storm objects after transferring to the downstream office since it was no longer their responsibility. In this experiment, upstream forecasters did not have any storm objects transferred to them. In NWS operations, it is reasonable to expect that most WFOs will experience an interval of increased workload as a WFO assumes responsibility for a storm that enters their CWA from an upstream office.

During the experiment, forecasters developed a strategy to manage this increased workload. Forecasters with an increased workload, particularly in the downstream WFO, delayed receiving a transfer of a storm object from an upstream WFO until their current workload diminished. Workload permitting, the downstream WFO would ask the upstream WFO to maintain ownership of the storm object, even if the object crossed into the downstream CWA.

4. Conclusions

In general, seamless transfers of ownership of storm objects occurred through direct and indirect methods of communication and other collaboration strategies. At times, increased workload and communication challenges hampered these transfers. The following sections offer a summary of results and a discussion of potential operational policies and software improvements to address unique collaborative situations and provide for a robust collaborative environment.

a. *Collaboration*

Aligned with the proposed FACETs paradigm shift to a continuous flow of hazard-related information, this experiment investigated collaboration between WFOs using storm objects to create seamless warnings across CWA boundaries. HS-PHI includes collaboration tools that enable the transfer of ownership from one WFO to another and domain permission requests. HS-PHI also allowed forecasters to view real-time edits to storm objects, plumes, and warnings from their adjacent WFOs. In addition to the collaboration tools, this experiment studied other communication strategies related to interoffice collaboration. Post-scenario and post-experiment human factors surveys and discussions yielded several interesting findings.

An analysis of themes and challenges showed that the forecasters organically developed a rudimentary concept of operations during the experiment. This concept of operations could potentially be used as a foundation for a future operational flow of warning operations. It can be summed up as a process of communication, handoff, and post-handoff.

Forecasters tended to rely on communication through chat and phone calls during the experiment, which was expected since they are accustomed to these methods in their current operations. Such communication was found to be important to ensure a quick, successful storm object transfer. This result agrees with related research that found a verbal handoff resulted in more successful handoffs and ensured transfer of accurate information (Arora and Johnson 2006). When communication broke down due to delays or non-communicative WFOs, transfers were not as fluid. During the experiment, forecasters stated that the frequency of communication increased, but unnecessary communication was reduced by using the collaboration tools in HS-PHI as a common reference.

While managing their own storm objects, forecasters in the downstream WFO were also situationally aware of objects from upstream WFOs and often prepared to communicate any necessary transfers if needed. Forecasters discussed transfers prior to initiating the transfer. Through viewing the same storm object in the HS-PHI software interface, forecasters from both WFOs were able to see details on the storm history, storm reports, warning methodologies, and PHI probabilities, thereby encouraging efficient communication.

The HS-PHI collaboration tools generally allowed for the quick and easy transfer of storm objects from one WFO to another. While most of the communication regarding storm objects occurred outside of the collaboration software, the collaboration tools permitted the forecasters to execute the transfer decision through either a "push" or "pull" operation to create seamless warnings across CWA borders. Determination on when to transfer objects between WFOs was based on many variables, including workload, a storm object's location relative to the border, societal concerns, and other considerations.

Mental workload also played an important role in forecaster collaboration. During periods of increased workload, object transfers were sometimes delayed. The delayed transfer of objects was not a procedural aspect of the experiment; rather, it was a natural development of forecaster collaboration. Forecasters used their own judgment to determine whether their workload was too high to receive a transfer. When workload issues were present, discussions occurred between forecasters for the upstream forecaster to maintain ownership of an object for a longer period of time, even if the object crossed into the downstream CWA.

After objects were transferred, the receiving forecasters often immediately updated the storm object attributes, including editing the object's duration or probability trend, or continuing or canceling an object or warning. Forecasters in the upstream WFO were sometimes concerned that the receiving forecaster could make significant changes to an object or warning that would lead to end users receiving inconsistent hazard information. These post-transfer changes could be due to numerous reasons, such as differences in warning philosophies or other societal concerns. Our research was not able to explore these reasons and the details of these outcomes. This experiment, however, provides opportunities to explore further human interactions and communication strategies within a high-fidelity warning environment.

b. *Proposed software functionality solutions*

Although the collaboration tools within HS-PHI provided the foundation for seamless object management and transfer by allowing a method to quickly transfer objects, forecasters suggested several improvements to software functionality to facilitate further collaboration.

1) Co-production

One popular suggestion among forecasters was the coproduction of storm objects and warnings, which would allow forecasters from multiple WFOs to interactively define, adjust, and collaborate on the production of a storm object within the HS-PHI software. The current format of the HS-PHI software only allows a neighboring WFO to *view* real-time edits to storm objects, but not participate in coediting the objects. In this coproduction state, forecasters may have the ability to communicate more effectively with their neighboring WFO by offering suggestions or edits in addition to collaboration.

2) Integrated communication

The addition of integrated chat functionality within the collaboration tools would be another feature to ensure effective interoffice collaboration and seamless transfers in the HS-PHI software. In the current version of the software, forecasters are able to accept or decline the transfer of an object but have no means to provide additional feedback on that decision without having to use another medium for communication. With integrated chat functionality, forecasters could collaborate and transfer objects within the same software.

3) Domain permission alternative

A proposed alternate method for domain permission suggested that permission be tied to the portions of plumes and warnings within separate CWAs, versus the storm objects. For example, the downstream WFO may elect to approve or deny the portion of the warning that extends into their CWA from an object owned by their neighboring WFO.

c. *Training and best practices*

The findings of this interoffice collaboration experiment could be used to inform the design of NWS policies and subsequent forecaster training and guidance regarding warning operations. Potential areas for policy and training based on this experiment's results include the following: when to transfer storm objects, when to maintain ownership of storm objects, when to force the transfer of an object (if necessary), when to request domain permission, and policies regarding consistent managing and messaging of storm objects across CWAs. This list is by no means comprehensive, as there are many other policy and training topics related to communication and collaboration not mentioned here. However, with the emergence of the collaborative forecast process (Uccellini and Ten Hoeve 2019) within the NWS, policies and best practices regarding interoffice communication and collaboration are being developed, and the development of applicable collaboration practices in the convective warning environment should be considered.

Within the scope of this project, potential best practices emerge. These include

1) INTEROFFICE WORKLOAD STRATEGY

Forecasters developed a strategy of maintaining ownership of objects, even beyond the scope of their CWA, if the workload of a receiving forecaster was too high. This flexibility allowed them to share the workload of managing storm objects. Careful communication was a significant aspect of successfully sharing the management of storm objects. The discussion of workload could even extend to multiple WFOs and occur before warning operations began.

2) COLLABORATION

With the goal of seamless hazard messaging across CWA borders, it is important to discuss differences in storm objects transferred from one WFO to another. Results from this experiment could be used to develop best practices and protocols for communication between WFOs to facilitate collaboration during warning operations. These best practices and protocols could include requirements for regular communication and policies regarding nonresponsive communication. Additionally, strategies for "pre-communication" could be developed in order for WFOs to anticipate and be prepared for object transfers. WFOs could develop thresholds for these object transfers that may be based on the meteorological situation or storm type, geographical boundaries, societal factors, or IDSS.

During this experiment, we observed that when neighboring WFOs had differing philosophies, forecasters would immediately update a transferred object. Policies likely need to be developed with regard to the immediate cancellation of storm objects to reduce inconsistencies in messaging. This could include a graduated approach to reducing objects or warnings as they enter another CWA. Additional software functionality could be included that would gradually reduce an object or warning over time, which would still provide consistent communication throughout the duration of the product and reduce erratic spatial changes to rapidly updating warning boundaries (otherwise known as the "windshield wiper effect"). Other software and best practice strategies should be considered to ensure continuity of severe weather hazard messaging across CWA borders before, during, and after events.

d. Importance of results

Although the impacts of interWFO collaboration on verification are not examined in this paper, future analysis and test bed experimentation can consider this. As Stumpf and Gerard (2021) show, TIM can improve average lead time and lead time equitability along the tracks of severe weather events. Robust collaboration between WFOs should improve performance, as the warnings will move forward steadily along with the storms without any border "roadblocks." A new gridded verification method has been developed that can be used to calculate verification metrics for TIM and test the above hypothesis (G. J. Stumpf and Stough 2022, unpublished manuscript). In addition to traditional verification measures, this new method can also calculate nontraditional measures, such as false alarm area to avoid overwarning areas outside the impacted locations, and departure time, which is useful for potential "all clear" products. This method can also be used to show how warning accuracy is affected by varying probabilistic warning thresholds using PHI. The verification method will be tested in future experiments.

A limitation to the current study is the limited sample size of human participants due to logistical challenges and with regard to funding WFO forecasters for a weeklong experiment. Each forecaster participated in numerous archived scenarios, while also participating in surveys and facilitator-led discussions. The resulting data gathered were comprehensive and provided numerous data points per participant. However, the data do not represent a statistically significant sample size and we are not presenting the data and conclusions as such; rather, this experiment should be regarded as a foundation for future research. Additionally, although the experimental design sought to simulate a realistic warning environment within NWS operations, it is not possible to simulate a high-fidelity warning environment in an experimental setting. There are many factors that cannot be artificially created, such as the real consequence of loss of life and property.

The HS-PHI software provided a new avenue for forecaster collaboration in which forecasters could coordinate the handoff and ownership transfer of plumes and warnings. Forecasters also received a glimpse into the mindset of the other WFO's forecasters based on the warning discussion and probability trends. This glimpse provided the forecasters opportunities to further collaborate on the objects during their phone calls or through chat. Although forecasters are comfortable using AWIPS II, chat, and phone, if all of these features and others are integrated into HS-PHI, it could become an all-in-one collaboration tool, with the possibility of other capabilities being included to foster this collaboration.

HS-PHI also blurs the lines of NWS CWAs, since objects can traverse borders without being clipped, and WFOs can own and control objects, plumes, and warnings of a neighboring CWA. In this regard, the HS-PHI software fosters collaboration between neighboring offices more than the current framework.

This experiment was designed to begin addressing the needed research questions in designing a collaboration tool for a dynamic, continuous PHI- and TIM-based severe weather warning system. The analysis of forecasters' performance and thoughts on interoffice collaboration provided valuable insights and feedback for future development and research. The processes developed during this experiment could lay the foundation for a future operational warning concept of operations. This experiment also shows the feasibility of HS-PHI software to facilitate the quick and easy handoff of storm objects supporting seamless warnings across CWA boundaries. Additional software revisions and experimental testing are needed to increase robustness and evaluate the performance of this concept of operations in challenging warning situations.

Acknowledgments. Funding was provided by NOAA/OAR under NOAA–University of Oklahoma Cooperative Agreement NA16OAR4320115 and NOAA Cooperative Agreement NA19OAR4320073, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of their institutions. The authors are appreciative of the contributions of many toward the development and completion of this experiment and research: Yujun Guo, Chris Golden, Tracy Hansen (ESRL/GSL), Tiffany Meyer, Justin Monroe, Travis Smith, Kristin Calhoun (NSSL), John Stoppkotte, Tom Bradshaw (NWS), and the NWS forecasters who participated in the experiment.

Data availability statement. The non-identifiable data collected in this study will be available upon request for two years.

REFERENCES

Argyle, E. M., J. J. Gourley, Z. L. Flamig, T. Hansen, and K. Manross, 2017: Toward a user-centered design of a weather forecasting decision-support tool. *Bull. Amer. Meteor. Soc.*, **98**, 373–382, https://doi.org/10.1175/BAMS-D-16-0031.1.

Arora, V., and J. Johnson, 2006: A model for building a standardized hand-off protocol. *Jt. Comm. J. Qual. Patient Saf.*, **32**, 646–655, https://doi.org/10.1016/S1553-7250(06)32084-3.

Beitlich, J. R., T. Funk, J. S. Schaumann, R. A. Wolf, A. W. Johnson, F. H. Glass, J. Stoppkotte, and R. A. Donavon, 2019: An overview of the National Weather Service central region tornado warning improvement program. *Seventh Symp. on Building a Weather-Ready Nation*, Phoenix, AZ, Amer. Meteor. Soc., 673, https://ams.confex.com/ams/2019Annual/webprogram/Paper354893.html.

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

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

Calhoun, K. M., 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**, 1–51, https://doi.org/10.1175/BAMS-D-21-0017.1.

FAA, 2015: Air traffic control, FAA order JO 7110.65W. Federal Aviation Administration, accessed 10 December 2015, https://www.faa.gov/regulations_policies/orders_notices/index.cfm/go/document.information/documentid/1028576.

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.

Hart, S. G., 2006: NASA-Task Load Index (NASA-TLX); 20 years later. *Proc. 50th Annual Meeting, Human Factors and Ergonomics Society*, San Francisco, CA, HFES, 904–908, https://doi.org/10.1177/154193120605000909.

——, and L. E. Staveland, 1988: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Adv. Psychol.*, **52**, 139–183, https://doi.org/10.1016/S0166-4115(08)62386-9.

Jackson, K., and P. Bazeley, 2019: *Qualitative Data Analysis with NVivo.* 3rd ed. Sage Publications, 376 pp.

James, J. J., 2021: Human factors and systems engineering analysis for development of partially automated severe weather warning methodologies. Ph.D. dissertation, University of Akron, 169 pp., https://etd.ohiolink.edu/.

——, and C. Ling, 2019: Forecaster's mental workload while issuing Probabilistic Hazard Information (PHI) during the 2018 hazard services PHI Hazardous Weather Testbed. *Ninth Conf. on Transition of Research to Operations*, Phoenix, AZ, Amer. Meteor. Soc., 9B.3, https://ams.confex.com/ams/2019Annual/webprogram/Paper355012.html.

——, ——, C. D. Karstens, J. Correia, K. Calhoun, T. Meyer, and D. Ladue, 2020: Forecasters' cognitive task analysis and mental workload analysis of issuing Probabilistic Hazard Information (PHI) during facets PHI prototype experiment. *Wea. Forecasting*, **35**, 1505–1521, https://doi.org/10.1175/WAF-D-19-0194.1.

Karstens, C. D., 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., G. J. Stumpf, A. V. Bates, J. LaDue, and A. E. Gerard, 2020: Interoffice collaboration: Current NWS practice and implications for a Probabilistic Hazard Information (PHI) future. *10th Conf. on Transition of Research to Operations,* Boston, MA, Amer. Meteor. Soc., 4.2, https://ams.confex.com/ams/2020Annual/meetingapp.cgi/Paper/369705.

——, and Coauthors, 2021: NWS severe warning philosophies across the CONUS: Implications for the implementation of FACETs. *11th Conf. on Transition of Research to Operations*, online, Amer. Meteor. Soc., 10.10, https://ams.confex.com/ams/101ANNUAL/meetingapp.cgi/Paper/384373.

Kuhlman, K. M., T. M. Smith, G. J. Stumpf, K. L. Ortega, and K. L. Manross, 2008: Experimental Probabilistic Hazard Information in practice: Results from the 2008 EWP spring program. *24th Conf. on Severe Local Storms*, Savannah, GA, Amer. Meteor. Soc., 8A.2, https://ams.confex.com/ams/pdfpapers/142027.pdf.

Landry, S. J., K. Levin, D. Rowe, and M. Nickelson, 2010: Enabling collaborative work across different communities of practice through boundary objects: Field studies in air traffic management. *Int. J. Hum. Comput. Interact.*, **26**, 75–93, https://doi.org/10.1080/10447310903025560.

Likert, R., 1932: A technique for the measurement of attitudes. *Arch. Psychol.*, **22** (140), 1–55.

NWS, 2020: WFO severe weather products specification. National Weather Service Instruction 10-511, 64 pp., http://www.nws.noaa.gov/directives/sym/pd01005011curr.pdf.

——, 2021: NWS county warning area map. Accessed 15 April 2021, https://www.weather.gov/gsp/nwsCWAMap.

Rantanen, E., 2004: Development and validation of objective performance and workload measures in air traffic control. Tech. Rep. AHFD-04-19/FAA-04-7, Federal Aviation Administration, Civil Aeromedical Institute, Oklahoma City, OK 130 pp., http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.132.1868.

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.

Sharples, S., A. Stedmon, G. Cox, A. Nicholls, T. Shuttleworth, and J. Wilson, 2007: Flightdeck and air traffic control collaboration evaluation (FACE): Evaluating aviation communication in the laboratory and field. *Appl. Ergon.*, **38**, 399–407, https://doi.org/10.1016/j.apergo.2007.01.012.

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

——, T. M. Smith, K. Manross, and D. L. Andra, 2008: The experimental warning program 2008 spring experiment at the NOAA Hazardous Weather Testbed. *24th Conf. on Severe Local Storms*, Savannah, GA, Amer. Meteor. Soc., 8A.1, https://ams.confex.com/ams/24SLS/techprogram/paper_141712.htm.

——, and Coauthors, 2018: Three years of hazard services—Probabilistic Hazard Information (HS-PHI) experiments at the NOAA Hazardous Weather Testbed. *29th Conf. on Severe Local Storms*, Stowe, VT, Amer. Meteor. Soc., 112, https://ams.confex.com/ams/29SLS/webprogram/Paper348359.html.

——, and Coauthors, 2020: Evolving hazard services Probabilistic Hazard Information for FACETs at the warning scale. *45th Annual Meeting*, online, National Weather Association, https://nwas.org/wp-content/uploads/2020/09/Agenda_Final.pdf.

Uccellini, L. W., and J. E. Ten Hoeve, 2019: Evolving the National Weather Service to build a weather-ready nation: Connecting observations, forecasts, and warnings to decision-makers through impact-based decision support services. *Bull. Amer. Meteor. Soc.*, **100**, 1923–1942, https://doi.org/10.1175/BAMS-D-18-0159.1.