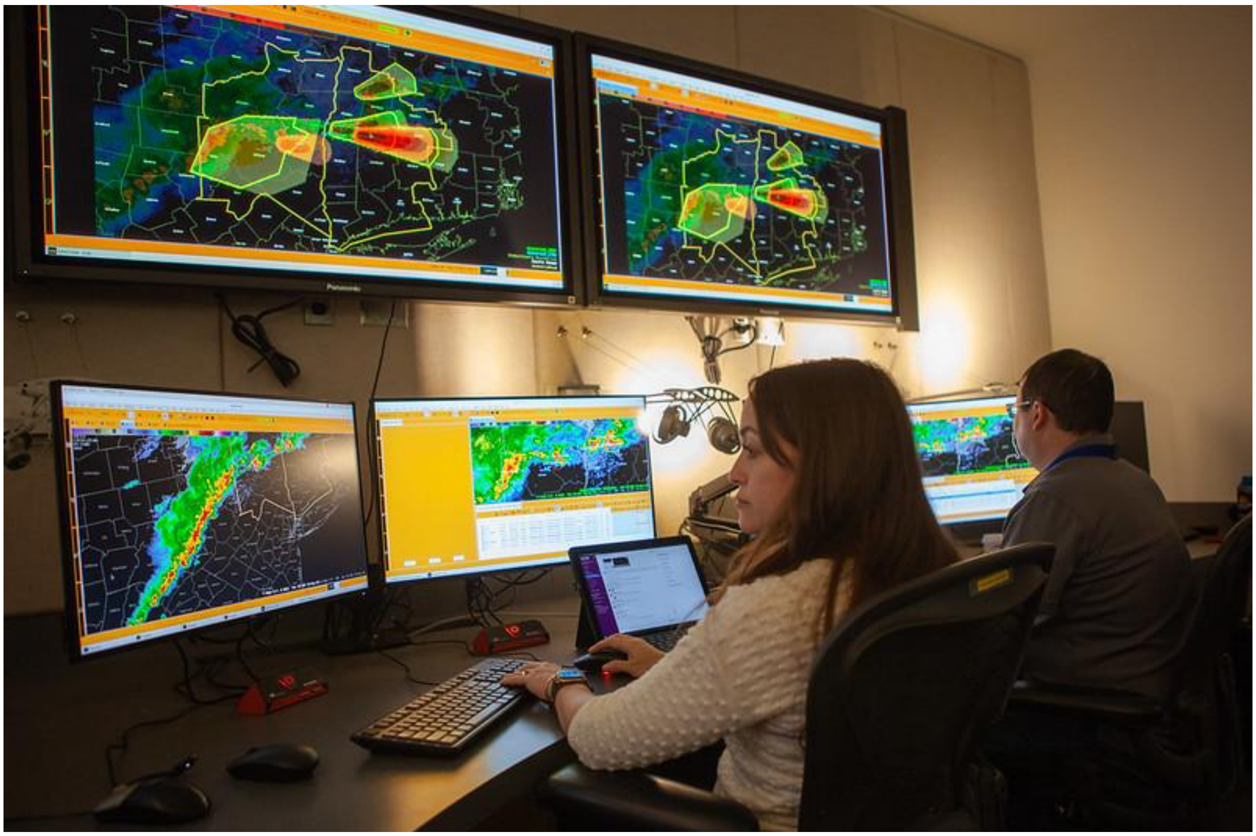
2020 EXPERIMENT SUMMARY

NOAA Hazardous Weather Testbed, Norman, OK

30 October 2020



Kodi Berry[1], M. Krocak[2,3], J. Ripberger[3], K. Klockow-McClain[2,1],
G. Stumpf[4,5], and J. James[6]

[1] NOAA/OAR/National Severe Storms Laboratory
[2] University of Oklahoma/Cooperative Institute for Mesoscale Meteorological Studies
[3] University of Oklahoma/Center for Risk and Crisis Management
[4] Colorado State University/Cooperative Institute for Research in the Atmosphere
[5] NOAA/NWS/Meteorological Development Laboratory
[6] University of Akron

TABLE OF CONTENTS

1. INTRODUCTION

The NOAA Hazardous Weather Testbed (HWT) is a joint project of the National Weather Service (NWS) and the National Severe Storms Laboratory (NSSL). The HWT provides a conceptual framework and a physical space to foster collaboration between research and operations to test and evaluate emerging technologies and science for NWS operations. The HWT emerged from the "Spring Program" which, for more than a decade, tested and evaluated new forecast models, techniques, and products to support NWS Storm Prediction Center (SPC) forecast operations. Now, the HWT consists of two primary programs: the original Spring Program, which is part of the Experimental Forecast Program (EFP), and the Experimental Warning Program (EWP).

Figure 1. A NWS forecaster using HS-PHI to manage a storm object and probability plume for an archived case.

The EWP tests and evaluates new applications, techniques, and products to support Weather Forecast Office (WFO) severe convective weather warning operations. This was the fourteenth year for warning activities in the HWT. Feedback was gathered from NWS operational meteorologists. The experiment participants issued experimental warnings, published live blogs, engaged in focus groups, and completed surveys. User comments were also collected during shifts, which have been used to inform product development. This input is vital to improving the NWS warning process, which ultimately leads to saved lives and property.

2. OVERVIEW

Due to the COVID-19 pandemic, all in-person HWT activities were suspended as of March 12, 2020. Most experiments were forced to postpone. The HWT EWP hosted two primary experiments over 4 calendar weeks to improve NWS severe weather warnings.

Cloud services were procured using NOAA's cloud contract vehicle starting on August 1, 2020. This contract allows for the investigation and implementation of remote HWT experiments using AWIPS in a cloud environment. With NOAA's inclusiveness of remote working employees and partners we feel that investigating methods to improve the opportunity and collaboration with as many NOAA employees and partners for the HWT is paramount.

EWP Experiment	Dates	Length	Number of Participants
Hazard Services - Probabilistic Hazard Information (HS-PHI) Interoffice Collaboration	Feb 10 - Feb 28	2 weeks	8 forecasters
Severe Weather & Society Dashboard	Sep 1 - Sep 10 Virtual	2 weeks	20 forecasters
PHI Prototype	Postponed until Spring 2021, Virtual		
Satellite Convective Applications	Postponed until Spring 2021, Virtual		
Radar Convective Applications	Postponed until Spring 2021, Virtual		
PHI End Users	Canceled		
Warn-on-Forecast	Postponed until Winter 2021, Virtual		
Hazard Services - PHI	Postponed until Winter 2021, Virtual		
Brief Vulnerability Overview Tool	Postponed until Summer 2021, Virtual		

Table 1. Details for the 2020 Experimental Warning Program.

3. PROJECT DETAILS AND RESULTS

Hazard Services – Probabilistic Hazard Information Interoffice Collaboration Experiment

Summary by Greg Stumpf and Joe James

Overview

NSSL developed a prototype tool for testing the early concepts of FACETs known as Probabilistic Hazard Information (PHI). This PHI Prototype Tool was evaluated by NWS forecasters and human factors experts in the HWT from 2014 to 2017. Several NOAA Joint Technology Transfer Initiative (JTTI) grants are funding an effort to transfer the PHI capabilities of the Prototype into AWIPS-II Hazard Services (HS), as well as develop additional functionality specific to NWS weather forecast office (WFO) environments.

The underlying concept behind PHI are storm objects that are created and periodically-modified by forecasters to delineate current areas of severe weather and project them 0-2 hours into the future. This is done using both probability plumes and warning polygons, each derived from the storm object information, and continuously-updating at one-minute intervals. The nature of this continuously-updating information requires independence from geo-political boundaries in order to provide a consistent and seamless transition of hazard and warning information between county warning areas (CWAs). This experiment was designed to test new HS-PHI functionality and best warning-scale practices to provide hazardous weather information and warning consistency across CWA boundaries.

The HS-PHI experiment was carried out for three weeks in late January into February 2020, with the first week being a "shakedown" of the system using two forecasters from the Norman, OK, WFO. The second two weeks included four pairs of participants from other WFOs nationwide, with three of the four pairs consisting of forecasters from the same home WFO. The experiment was set up to simulate operations in two separate WFOs, each with two NWS forecaster participants, creating PHI and warnings for their respective CWA, and collaborating with each other to handle storms near the boundaries of their adjacent CWAs. This was accomplished by separating the two teams of forecasters in two locations at the National Weather Center: 1) the Hazardous Weather Testbed operations area, and 2) the NSSL Development Laboratory. Each location was equipped with AWIPS workstations that were on the same network. The forecaster pairs remained sequestered from each other during the exercises, and collaboration between the WFOs was made possible by a chat room and telephones.

Two new HS-PHI components were tested during the experiment. The first was an "ownership transfer" function, which allowed one WFO to transfer the ownership of a storm object to their neighboring WFO as the storm approached the boundary between their two CWAs. The second was a "domain permission" capability which would allow one WFO to create a new storm object within the domain of their neighbor's CWA. This was typically done with storms just outside the boundary of their CWA that were projected to enter into their CWA soon. The HS-PHI software also provided a live view of the storm objects that were being edited by the other WFO.

Archived data events were used to perform the evaluations; there were seven cases from different locations nationwide, that each had unique adjacent CWA domains and represented a variety of different severe storm types (e.g., squall lines, long-tracked tornadoes, etc.). The study aimed to explore several ideas to represent realistic challenges currently faced in the collaboration between WFOs. These ideas include priming the teams individually to manipulate their perception of their neighboring offices and implanting different cultures with each team by giving them different directives on local warning policies. In addition, fictitious impact-based decision support services (IDSS) "events", such as outdoor festivals, located near the shared border of the two CWAs, were strategically inserted to challenge the forecasters' decision making process as storms crossed CWA boundaries. The forecasters were directed to collaborate as they saw fit during these scenarios.

As with previous experiments, this evaluation included human factor experts who recorded video and audio, and administered surveys and interviews to measure mental workload, confidence, effectiveness of HS-PHI collaboration, and software usability. Several suggestions for improved software functionality were offered by the forecasters, some of which will be incorporated in order to make the software more robust.

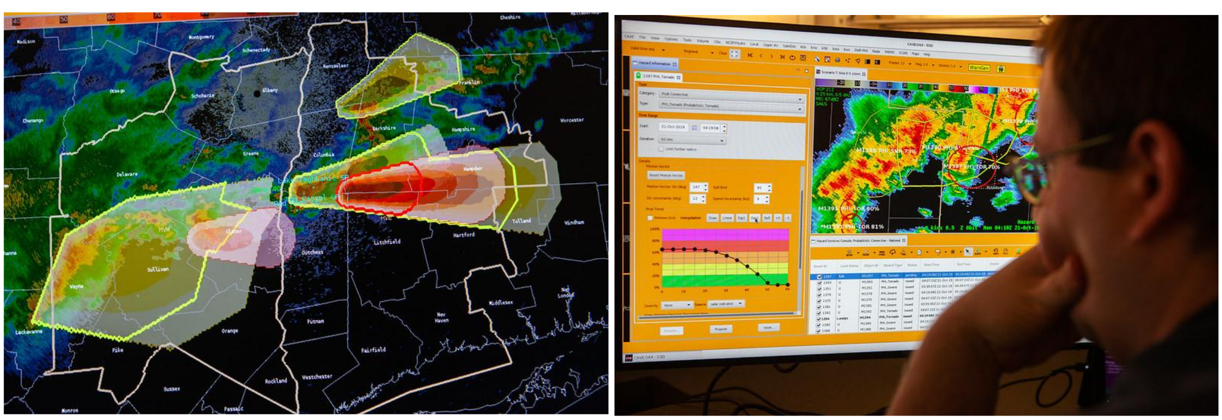


Figure 2. Images from the HS-PHI experiment. Left: HS-PHI output from a scenario in upstate New York - the yellow (red) shading show hail/wind (tornado) probability plumes; the yellow (red) polygons show severe thunderstorm (tornado) warnings; the large white polygons are two county warning areas (CWA). Right: Forecaster using HS-PHI on an AWIPS-II workstation.

2020 Accomplishments

The following is a list of accomplishments from the 2020 HS-PHI Interoffice Collaboration Experiment:

- Completed the development functionality to facilitate the seamless transition of PHI and warnings across adjacent CWA borders:
  - Ownership transfer.
  - Domain Permission.
- Developed seven archived cases from 2012-2019 to test the software on a variety of severe weather conditions. Cases included:
  - Two training cases, and one operations case, with isolated severe storms moving across a CWA border.
  - Multiple severe storms moving along a non-straight CWA border, resulting in some storms crisscrossing the border.
  - Single isolated violent long-tracked tornado event.
  - Isolated storms and a squall line crossing a CWA border.
  - Multiple storms moving across a CWA border over the course of 5 hours
- Shakedown test of HS-PHI in the HWT from 27-31 January 2020 with 2 NWS Norman forecaster participants.
- Operations of HS-PHI in the HWT for two alternating weeks from 10-14 February 2020 and 24-28 February 2020 with 4 NWS forecaster participants per week. The outcomes of this test included:
  - Gathering feedback on software performance and design, with bug-fixes and improvements developed and tested during the off-weeks of the test.
  - Collecting forecaster workload data in collaboration with human factors scientists.
  - Capturing discussions on the FACETs and PHI concepts in NWS severe weather warning operations, concentrating on how adjacent forecast offices would collaborate and share severe storm objects to provide seamless service across forecast area boundaries.
- Key Takeaways:
  - Forecasters thought this kind of workflow with collaboration was possible (assuming they had adequate staffing and suggested software improvements added).
  - Add a co-production capability, so that two forecasters from two WFOs can create and edit storm objects as a team, versus one forecaster doing the editing while the other forecaster observes.
  - Domain permission should not just be based on a storm object, but rather the plumes and warnings. Perhaps each WFO "authorizes" the portions of the plumes and warnings in their CWA.
  - Add chat capabilities directly into the collaboration pop-up windows.
  - Add filters to dim or hide objects not owned by your WFO.
  - Collaboration allows WFOs with lighter workloads to help out a neighboring WFO which is experiencing high workload to manage storm objects near the border.
  - Additional social science work is needed to overcome differences in culture and warning decision approaches between WFOs.

Human Factors Analysis

a. Mental Workload (NASA TLX) Survey

The NASA-TLX survey is a questionnaire-based mental workload rating tool. The tool encompasses six sub-dimensions of mental workload: mental demand, physical demand, temporal demand, performance, effort, and frustration. The analysis of mental workload includes a weighting dimension used to calculate an overall workload score. The raw scores of the mental workload ranges from 1 to 100, where 1 represents extremely low workload and 100 represents extremely high workload. The ratings were averaged from all of the cases for each of the six sub-dimensions of workload, and the importance factors were calculated for each aspect.

Figure 3 shows the average mental workload for all forecasters from all archived hazardous weather events. Each bar in the figure represents the average mental workload for each of the six sub-dimensions of workload. The red line represents the average workload. The average workload for 2020 HS-PHI was 58.7 (out of 100, standard deviation 19.9), which is an increase from the 2019 workload average of 54.4. All workload sub-dimensions increased from 2019.

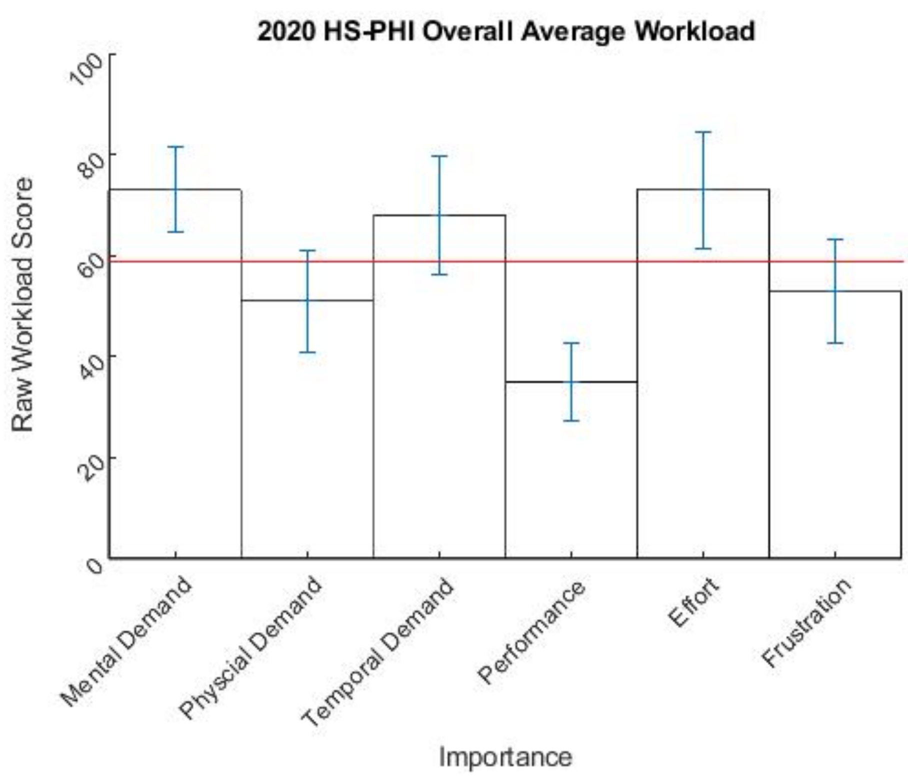


Figure 3. NASA-TLX workload sub-dimensions with the average mental workload for the 2020 HS-PHI Collaboration experiment represented by the red line.

Figure 4 shows the average mental workload for the past four years of the HS-PHI experiment. The 5-year average of HS-PHI workload is 53.8.

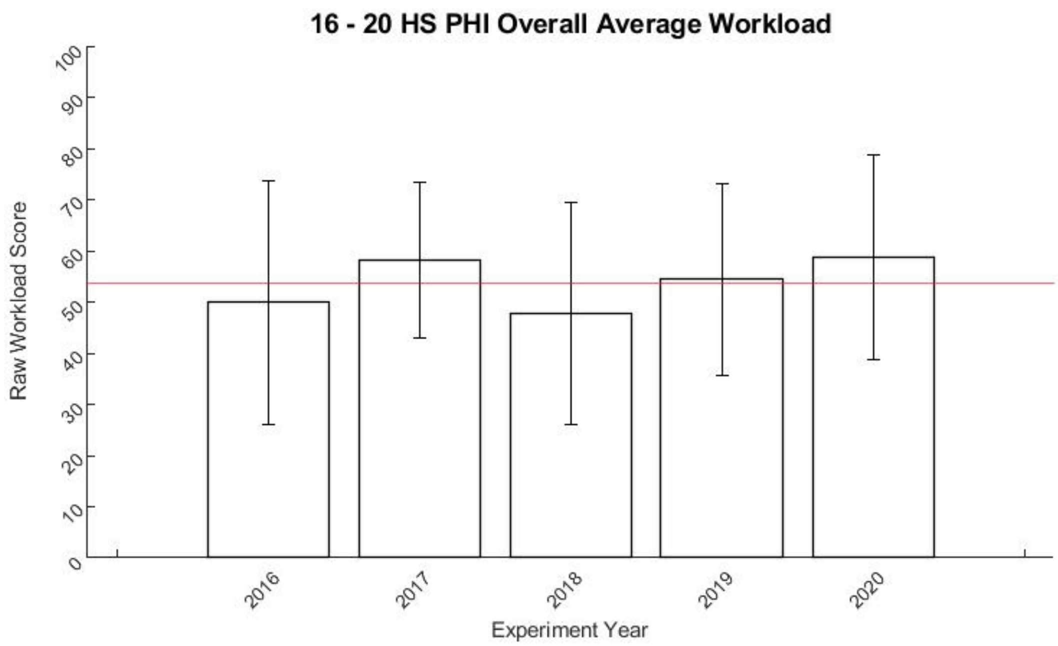


Figure 4. NASA-TLX mental workload with the average mental workload for all five years of the HS-PHI experiment represented by the red line.

The downstream CWAs consistently reported higher mental workload of 10 points or more across scenarios 3, 5, 6, and 7 (scenario 4 was a border case without a clear upstream/downstream CWA), Fig. 5. Upstream CWAs reported an average mental workload of 52 and the downstream CWAs reported an average mental workload of 65.5 out of 100. Scenarios 3 - 7 were presented consecutively during each week of the experiment, there appears to be a slight workload decrease in each consecutive scenario, until scenario 7, which was an all-day scenario, compared to 2.5-hour events for scenarios 3, 4, 5, and 6.

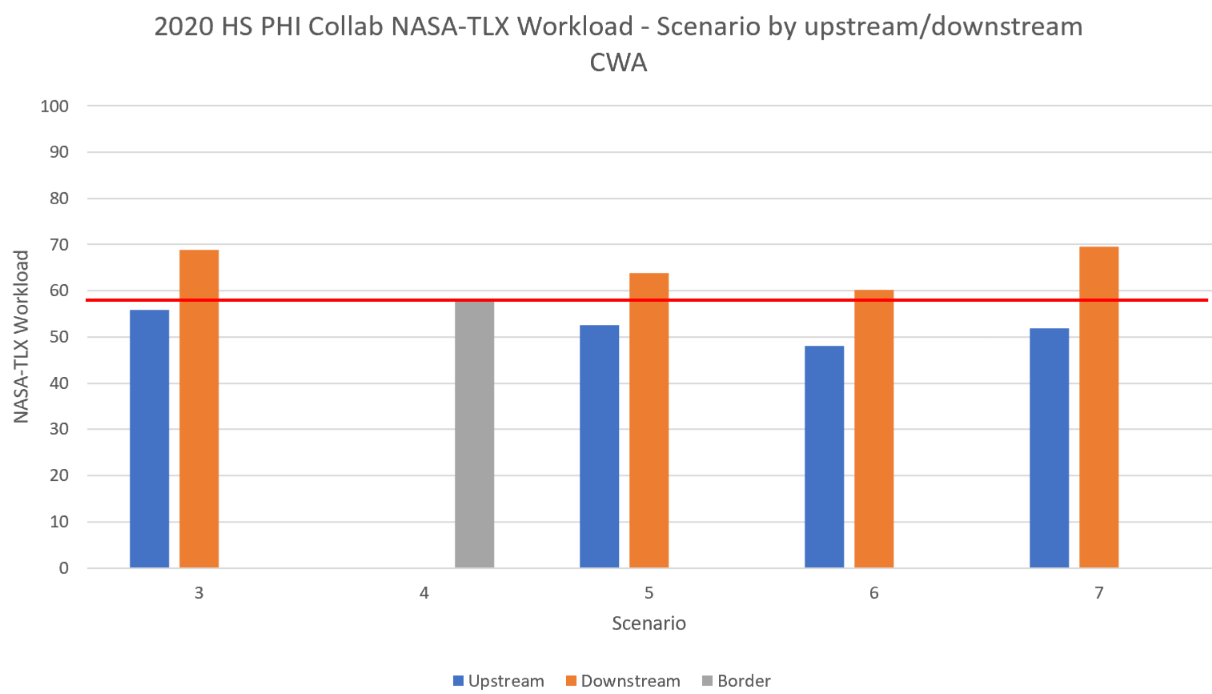


Figure 5. Forecaster NASA-TLX workload comparing upstream and downstream WFOs for each scenario in the 2020 HS-PHI Collaboration experiment.

b. Collaboration Survey

The collaboration survey was developed to understand how forecasters feel about the collaboration tools in the Hazard Services tool. Two different surveys were developed, one for end of week and one for end of case. Both surveys are a multiple-choice answer on a 1 - 7 Likert scale. The end of case survey was given to show the variation between different case types and situations, the end of week survey was given to gain an overall understanding of forecasters' experience with collaboration during the whole experiment.

Collaboration was very similar considering upstream and downstream CWAs. Additionally, comparing across scenarios, the confidence showed little variation. Confidence did not increase as the week of experiments progressed, but stayed consistent.

Forecasters rated the end-of-week collaboration survey a 5.0 out of 7 (slightly agree). Highest agreement was for statements, “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 week 2 forecaster 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”.

Forecasters found the least agreement 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)". Regarding fostering collaboration, a forecaster commented, "I could see what the neighbors thought process was, but the actual collaboration was done outside of the AWIPS2 ecosystem".

2020 HS PHI End-of-Week Collab Survey	
With the HS-PHI software, I felt more aware of the neighboring WFO's warning decision process.	*4.8*
Tracking hazards using 2D objects improved my ability to collaborate with my neighboring WFOs.	*5.0*
The ownership transfer tool and the domain permission tool improved my ability to collaborate with my neighboring WFOs.	*5.3*
The HS-PHI software fostered more effective collaboration between me and my neighboring WFO.	*5.1*
The HS-PHI software fostered more effective collaboration between me and my neighboring WFO when our warning and/or PHI philosophies differed.	*4.3*
The HS-PHI software fostered more efficient collaboration warnings.	*5.1*
I can develop better strategies for collaboration when issuing PHI plumes and warnings using the HS-PHI software.	*5.3*
Overall Average	***5.0***

Table 2. End of week collaboration confidence survey (7-point scale)

End-of-Case Collaboration Question	***Mean***
1. With the HS-PHI software, I felt aware of the neighboring WFO's warning decision process.	*4.9*
2. I felt comfortable when a PHI plume existed in my CWA and the neighboring CWA at the same time.	*5.9*
3. I felt comfortable when my neighboring office had control of a PHI plume that was also in my CWA.	*5.4*
4. I felt comfortable when my neighboring office made a decision that	*5.3*

modified the PHI plume within my CWA.	
Decisions regarding PHI plume average	***5.5***
5. I felt comfortable when a warning existed in my CWA and the neighboring CWA at the same time.	*5.8*
6. I felt comfortable when my neighboring office had control of a warning that was also in my CWA.	*5.2*
7. I felt comfortable when my neighboring office made a decision that modified the warning within my CWA.	*5.0*
Decisions regarding warning average	***5.3***
8. Tracking hazards using 2D objects helped me to collaborate with my neighboring WFOs.	*5.4*
9. The ownership transfer tool and the domain permission tool helped me to collaborate with my neighboring WFOs.	*5.3*
10. The HS-PHI software fostered effective collaboration between me and my neighboring WFO.	*5.7*
11. The HS-PHI software fostered effective collaboration between me and my neighboring WFO when our warning and/or PHI philosophies differed.	*4.6*
12. The HS-PHI software fostered efficient collaboration on PHI plumes and warnings.	*5.3*
13. I easily developed strategies for collaboration when issuing PHI plumes and warnings using the HS-PHI software.	*5.9*
Total average	***5.4***

Table 3. End of case collaboration survey results

Plans for Future Experiments

The following is a list of goals for the next iterations of the HS-PHI experiments:

- Set up a software development and testing environment using cloud computing.
- Continue software refinement and development of new functionality, including:
  - Improvements to Threats-In-Motion (TIM) and warning product generation.
  - Addition of a past trend widget.
  - Addressing performance issues and software bugs.

- A virtual HWT experiment during Winter 2021, including a greater emphasis on TIM.
- Integration of HS-PHI into the operational version of Hazard Services.
- Complete development of FY17 JTTI version of HS-PHI by 30 Sep 2021 to RL7.

Web Presence

vlab.ncep.noaa.gov/group/hazardous-weather-testbed/hwt-2020-hs-phi-collaboration-resources

Project Contacts

Alyssa Bates	CIMMS/NWS-WDTD	Project Lead PI
Greg Stumpf	CIRA/NWS-MDL	HWT Experiment Coordinator / Project Co-PI
Kim Klockow-McClain	CIMMS/NSSL	Project Co-PI
Darrel Kingfield	GSL (formerly CIRES/GSL)	Project Co-PI
Chen Ling	U. Akron	Project Co-PI
Alan Gerard	NSSL	Project Co-PI
James LaDue	NWS-WDTD	Project Co-PI
Kevin Manross	CIRA/GSL	Lead Developer
Chris Golden	CIRES/GSL	Developer
Yujun Guo	CIRA/GSL	Developer
Joe James	U. Akron	Project Scientist
Pat Hyland	CIMMS/NSSL	Project Scientist
Tiffany Meyer	CIMMS/NSSL	AWIPS-II Support
Justin Monroe	CIMMS/NSSL	AWIPS-II Support

Severe Weather and Society Dashboard Experiment

Summary by Makenzie Krocak, Joe Ripberger, and Kim Klockow-McClain

Overview

The 2020 Severe Weather and Society Dashboard experiment was conducted from September 1 through September 10, 2020. The major goal of this experiment was to assess the utility of a new type of product with forecasters - one that provides social and behavioral data about the communities that NWS Forecasters serve. Containing data from the annual Severe Weather and Society Survey, the Dashboard (WxDash) is the first primarily social science tool to be tested in the HWT. A wide variety of data was presented to participants, from composite indices that measure tornado warning reception, understanding, and response, to geographic risk perceptions of different weather hazards and demographic variables. The overall utility of the dashboard tool was evaluated along with the usefulness of the data within the dashboard and the timescales on which the data would be most useful.

Experiment Details and Results

During the experiment, participants were asked to complete a few different activities. On day one, forecasters first discussed the goals of their communication strategies across a range of timescales: from general outreach before a specific event is being forecasted to the warning time frame. Then, participants took a pretest where they evaluated how they thought residents in their forecast area answered questions about severe weather products. This activity was included to get forecasters thinking about their populations and their own preconceived notions about them. Then researchers introduced the dashboard and asked participants to use the dashboard to find the answers to the pretest. This activity was particularly helpful because participants were given the opportunity to explore the dashboard on their own and evaluate their knowledge of the populations they serve. Many participants thought that their original thoughts about the population were too pessimistic, that many people understand more about the weather than they originally thought. However, participants also expressed disbelief at some of the results, particularly about individuals' self-rated abilities to read maps.

Participants also had the opportunity to create communication products during different event time scales. During the general outreach portion, many forecasters used the dashboard to inform the content of the product they were creating. For example, the population in central Iowa tends to have a relatively high understanding of severe weather products. Therefore, forecasters felt comfortable focusing on more advanced topics like the timescales of watches and warnings instead of just the general definitions. Other areas, like central South Carolina, have a high proportion of the population living in mobile housing, prompting forecasters to focus on mobile home safety in their communication products.

Figure 6: A screen capture of participants in the Severe Weather and Society Dashboard experiment held virtually in September, 2020.

Over the course of the experiment, some common themes emerged about how the dashboard would be useful, when it would be useful, and future ideas to improve it. First and foremost, all of the respondents thought that the dashboard provides useful data. There was widespread agreement that the dashboard would be useful during forecaster training to visualize population characteristics and to highlight communication strategies that should be focused on. Many of the participants also said that this tool could be useful to describe quantitatively how social and behavioral data is useful to their jobs as forecasters. There was also agreement among participants that this tool would be most useful at longer timescales. By the time forecasters are in warning mode, they are managing too much data to be constantly looking at the dashboard for population data. Similarly, participants suggested a training module or course that specifically focuses on interpretation and use of social and behavioral data, noting that the dashboard is a great tool to have available where all of this data is kept.

While there was widespread agreement that the dashboard is a useful tool, there were many suggestions to help improve its utility for forecasters and their daily workflow. First, a complete explanation of the statistics used to display the data is necessary for ease of interpretation. The current display features a few different percentile measures that are confusing to interpret on the fly. Similarly, forecasters suggested a "fact sheet" for each CWA that can be referenced easily on the front page of the dashboard. This would allow forecasters to get an overview of their area and then prompt them to dig into the data when time permits. Another visualization suggestion was to incorporate this data into AWIPS, which would allow forecasters easier access to the data when they are in more information-intensive scenarios (like the warning time frame).

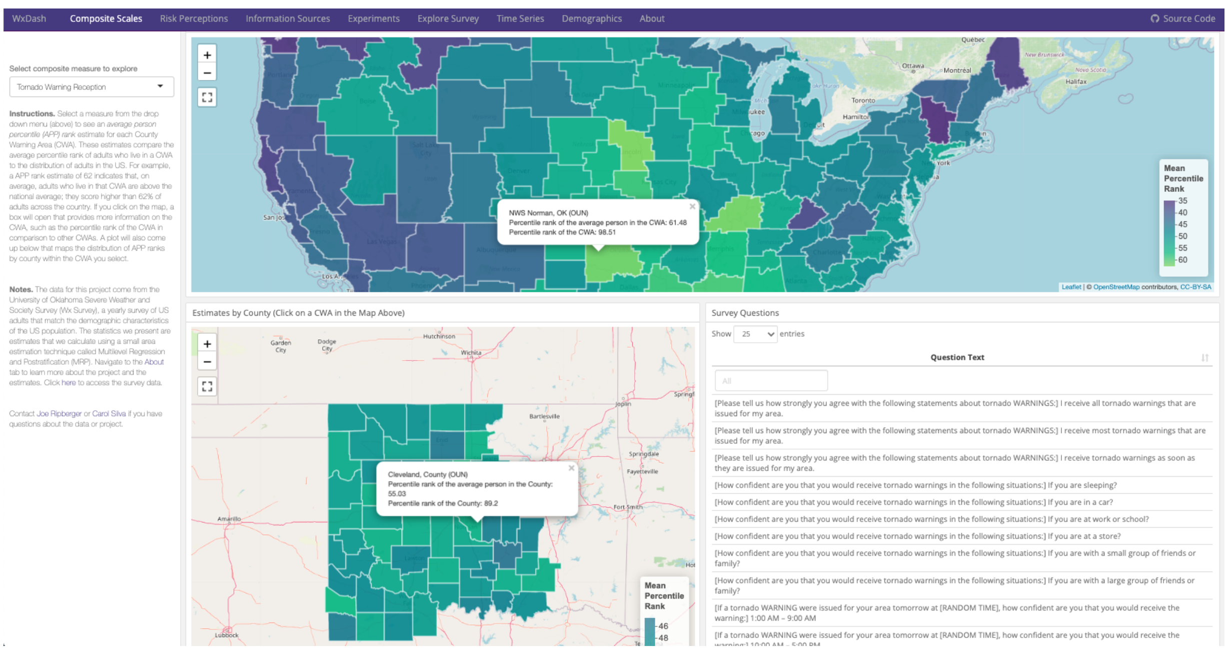


Figure 7: A screen capture of the prototype Severe Weather and Society Dashboard. Data comes from the Severe Weather and Society Survey. A description of the 2020 iteration can be found at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/EWOCUA.

In addition to the myriad of suggestions about data visualization and explanation, participants also had suggestions for future research avenues. Many of them want to see this data tested with local emergency managers to get their input on the validity and value of a social science tool. They would also like to see new datasets incorporated into the dashboard, like typical working conditions, working hours, home structures (like basements), and floodplain areas. All of these datasets would give forecasters a better idea of the hazards and timeframes that they should focus on when communicating the weather and the potential impacts of hazardous events. The dashboard experiment provided validation that social and behavioral data is useful for forecasters, and more importantly raised new questions and suggestions for ways to incorporate this data into forecasters' daily workflow. Future experimentation will focus on understanding how partners could use the dashboard data and how changes to the dashboard may allow community-level data to fit into the current forecaster workflow.

Future Plans

In addition to future experiments that specifically test the utility of the severe weather dashboard, researchers are pursuing paths for operationalization of the dashboard and future community-level data. Surveys focusing on other weather hazards (like tropical cyclones, winter storms, and fire weather) are currently under development, with future plans to create dashboards similar to the severe weather prototype. Researchers welcome feedback from forecasters regarding if and how this data may help guide risk communication practices in different offices.

Web Presence

Severe Weather and Society Dashboard https://crcm.shinyapps.io/WxDash/

Project Contacts

Joe Ripberger	CRCM	POC and PI
Kim Klockow	CIMMS/NSSL	Co-PI and Research Scientist
Makenzie Krocak	CRCM/CIMMS	Research Scientist
Jinan Allan	CRCM	Graduate Researcher
Sean Ernst	CRCM	Graduate Researcher

4. PERSONNEL

Alan Gerard
Acting HWT Co-Executive Officer alan.e.gerard@noaa.gov

Kodi Berry
Acting HWT Co-Executive Officer kodi.berry@noaa.gov

Kristin Calhoun
Acting HWT Co-Executive Officer kristin.calhoun@noaa.gov

Justin Monroe
EWP Technical Lead justin.monroe@noaa.gov

Jonny Madden
EWP Technical Advisor jonathan.madden@noaa.gov

5. ACKNOWLEDGMENTS

We would like to acknowledge the contributions of the following individuals: Tanya Riley, Tracy Reinke, Keli Pirtle, Emily Summars, James Murnan, Karen Cooper, Carrie Langston, and Vicki Farmer.

HWT Infrastructure Support was provided, in part, by the U.S. Weather Research Program and Joint Technology Transfer Initiative. Research was funded by the U.S. Weather Research Program, Joint Technology Transfer Initiative, National Severe Storms Laboratory, and NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA16OAR4320115, U.S. Department of Commerce.