

Exploring Applications of Eye Tracking in Operational Meteorology Research

BY KATIE A. WILSON, PAMELA L. HEINSELMAN, AND ZIHO KANG

BACKGROUND. Recently, eye tracking has been used within the meteorology community to assess communications of weather information to the public. Drost et al. (2015) used eye tracking to study the impact of a weathercaster's gesturing during a televised weather forecast on viewers' attention. Their analysis revealed that while gesturing did not impact viewers' retention of information, it did redirect viewers' attention to different elements on the screen. Eye tracking was also used by Sherman-Morris et al. (2015) to investigate the impact of different legend colors and content in hurricane storm surge graphics on participants' ability to accurately interpret threat levels. Although significant differences in accuracy were not found across legends of different color and content, participants' eye-tracking data indicated they struggled most when the legend color was a shade of blue and the values were in feet. Studies such as these are helping the United States work toward becoming a Weather Ready Nation. A Weather Ready Nation is one that builds community resilience to increasing vulnerability of extreme weather and water events (NOAA 2015). Lindell and Brooks (2013) summarized a number of major issues that a Weather Ready Nation workshop in 2012 identified as requiring attention, and conducting cognitive

research in laboratory experiments to understand users' interpretations of forecasts and warnings was one issue they identified. The studies described by Drost et al. (2015) and Sherman-Morris et al. (2015) demonstrate ways in which eye tracking is being used to help address this issue.

Another issue identified in the 2012 Weather Ready Nation workshop was the need to study forecasters through behavioral research (Lindell and Brooks 2013). Highlighted was the need for research to develop an understanding of forecasters' decision-making processes and how they differ between individuals and the National Weather Service (NWS) regions. To date, forecaster decision-making processes have been examined using a variety of qualitative methods. For example, an ethnographic approach was used by Daipha (2015) to observe and study how forecasters collect and use information in the forecast office. Root Cause Analysis is also performed by forecasters after an event has occurred so that they can evaluate their own warning decisions (Quoetone et al. 2009). Root Cause Analysis encourages forecasters to reflect on their decision-making processes and helps uncover reasons for why problems occur. The Critical Incident Technique has also been used in research to gather stories of forecasters' descriptions of past events and what their associated behaviors were (LaDue et al. 2010). Furthermore, research in the National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed has used surveys and blogs to collect forecasters' feedback of new products tested during warning operations (Calhoun et al. 2014). A retrospective recall method has also been used in the Hazardous Weather Testbed to study individual forecaster's cognition associated with radar data interrogation (Heinselmann et al. 2015; Bowden et al. 2015). This method collects video-cued recall information while forecasters watch a playback video of their on-screen activity and verbalize their

AFFILIATIONS: WILSON—Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma; HEINSELMAN—NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma; KANG—School of Industrial and Systems Engineering, University of Oklahoma, Norman, Oklahoma

CORRESPONDING AUTHOR: Katie Wilson, 120 David L. Boren Blvd, Norman, OK 73072

E-mail: katie.bowden@noaa.gov

DOI:10.1175/BAMS-D-15-00148.1

©2016 American Meteorological Society

past thought processes. Specifically, this method yields detailed information about what forecasters see, think, and do while interrogating radar data. Although retrospective recall data have provided incredible insight, the complexity of forecasters' decision processes means that the use of qualitative methods alone does not fully capture the intricate cognitive processes of forecasters.

To our knowledge, eye tracking has not been applied to study NWS forecasters' decision-making and related cognitive processes. However, applications of eye tracking in a variety of research domains, including the studies carried out by Drost et al. (2015) and Sherman-Morris et al. (2015), suggest that this tool could enrich our understanding of how forecasters use information to make decisions. Studies in research domains such as air traffic control and medicine demonstrate how eye tracking can be used to ask questions that—in an analogous sense—we may wish to answer in operational meteorology. For example, Kang and Landry (2014) used eye tracking to analyze how novice and expert air traffic controllers' eyes scanned a radar display during aircraft conflict detection tasks. Kang and Landry (2014) found that training novices with experts' scanpaths reduced the novices' number of false alarms. We may wonder in operational meteorology how low- and high-performing forecasters' scanpaths of weather radar data differ, and whether such information may be helpful during training. Wood et al.'s (2013) study on visual expertise of radiologists during detection and diagnosis of skeletal fractures is also relatable to operational meteorology. After all, forecasters use radar data to detect the potential for severe weather and then correctly diagnose what type of threat they expect. In Wood et al.'s (2013) study, radiologists' eye gaze data were used to measure their accuracy and speed, which are also measures used to analyze forecaster performance.

EXAMPLE: UNDERSTANDING A FORECASTER'S DECISION PROCESS. To explore how forecasters' eye gaze data may enrich our current understanding of their decision processes, we collected an NWS forecaster's eye gaze data as he interrogated radar data from one weather event, and subsequently obtained his retrospective recall. Eye-tracking research is built on the foundation of the eye-mind hypothesis, such that we assume a person's eye gaze indicates where their attention is and what is at the "top of the stack" of their cognitive processes



FIG. 1. Forecaster interrogating 1-min base velocity and reflectivity phased array radar data using the Warning Decision Support System-Integrated Information. The Tobii TX300 eye tracker is positioned below the monitor.

(Just and Carpenter 1976). Therefore, measuring forecasters' eye gaze behavior may provide a way for us to learn about their cognition at a deeper level. The goal of this short study was not to draw conclusions about forecaster cognitive processes, but to think about what type of information eye-tracking methods can provide for learning about cognitive processes that our current qualitative methods do not.

During this short study, the forecaster viewed a 39-min-long severe hail and wind event from 16 July 2009 in displaced real time and was asked to make warning decisions as he saw necessary. During this event, a nonsevere northern storm and severe southern storm moved south toward Oklahoma City, Oklahoma. The nonsevere northern storm was well developed at the beginning of the case, while the southern storm was captured from early in its initiation. The forecaster viewed 1-min base velocity and reflectivity phased array radar updates (Zrnić et al. 2007 and Heinselman and Torres 2011) using the Warning Decision Support System-Integrated Information (WDSS-II; Fig. 1). The forecaster was able to loop through radar data, navigate in time and by elevation using function keys, and zoom in and out. Warnings were issued using a polygon tool located in the control panel.

Throughout the simulation, the forecaster's eye gaze data were collected using the Tobii TX300 eye-tracking system (Fig. 1). This system sat below the forecaster's computer monitor, from which an infrared camera detected the location of his pupils and corresponding eye movement on the screen. We viewed the forecaster's eye gaze data using the Tobii

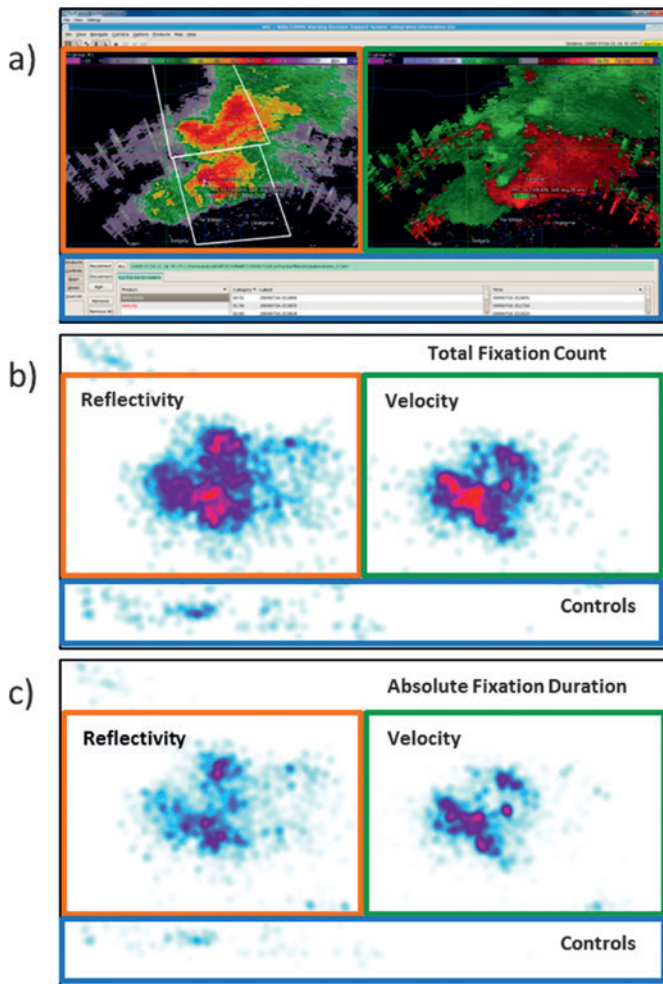


FIG. 2. (a) The Warning Decision Support System-Integrated Information display divided into three areas of interest: reflectivity (left panel, orange box), velocity (right panel, green box), and controls (bottom panel, blue box). Heatmaps were created for the (b) total fixation count and (c) absolute fixation duration for the entire case. Within the heatmaps, red values indicate a higher fixation count and absolute fixation duration, and blue colors indicate a lower fixation count and shorter absolute fixation duration.

Studio 3.3.0 software, and used a velocity-threshold filter algorithm to identify when and where the forecaster's eye fixations occurred (Olsen 2012). The forecaster's fixations describe times when his eye gaze momentarily focused on a specific location. The focus is long enough such that he was able to encode and process information (Poole and Ball 2006). The fixation algorithm provided timestamp, duration, and x and y position information for each fixation that the forecaster made. Additionally, we were

able to see whether his fixations were made within the reflectivity, velocity, or control panels by creating three separate areas of interest (AOIs; Fig. 2a). Defining AOIs in eye-tracking analysis is common practice, as this method allows for different types of information presented on the same screen to be distinguished from one another. While the reflectivity and velocity panels presented information about the storms, the control panel provided a polygon tool for issuing warnings.

We looked at two measures of fixation during this study: fixation count and fixation duration. Higher numbers of fixation count on a particular AOI indicate that the information was either more noticeable or important, whereas longer durations of fixations on a particular AOI indicate that the information was either more engaging or that a greater mental effort was required to extract the information (Poole and Ball 2006). Unlike retrospective recall information, the forecaster's eye gaze data can be used to obtain detailed information about the spatial distribution and temporal trends of these fixation measures in each of the three AOIs. We were interested to see how these fixation measures compared across the three AOIs for the full simulation and how their values changed as the weather scenario evolved. Additionally, we looked at how the forecaster's fixation measures corresponded to the information provided in his retrospective recall, and whether together these two datasets offer a more holistic and accurate understanding of his decision process.

COUNTS AND DURATIONS OF EYE FIXATIONS.

Heatmaps are visualizations of the overall spatial distribution of eye fixations within specified AOIs (Fig. 2). In Fig. 2b, we see that the forecaster fixated most often on the reflectivity and velocity AOIs, and least often within the controls AOI, indicating focus on data interrogation and limited use of the control panel to issue warning polygons (Figs. 2b and 2c). The distributions of 1-min fixation count and mean duration support this interpretation (Fig. 3). Applying the Wilcoxon rank sum test, statistical significance ($p < 0.05$) was established for the difference in median values of 1-min fixation counts and 1-min mean fixation durations across all three AOIs (Figs. 3a and 3b). Variations in the spatial patterns of total fixation count seen in the reflectivity and velocity AOIs suggest the forecaster interrogated these fields differently. A comparison of these heatmaps to the most typical positioning of radar data on

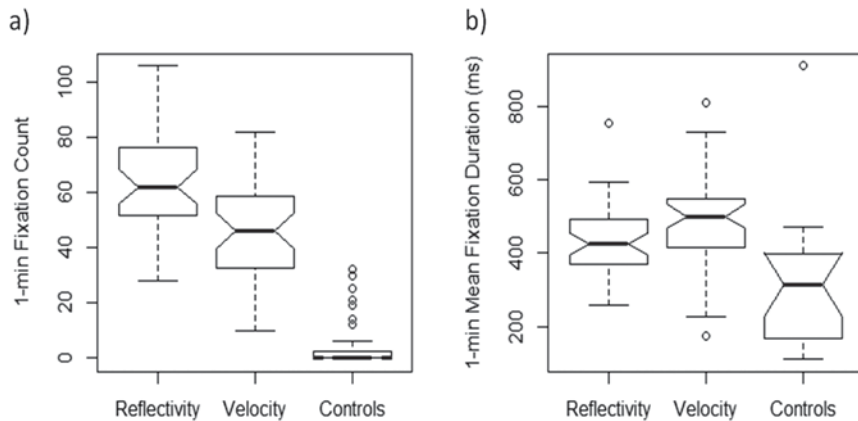


FIG. 3. Boxplots showing the distribution of (a) the 1-min fixation count and (b) the 1-min mean fixation duration for the reflectivity, velocity, and controls AOIs. Boxplot whiskers indicate minimum and maximum values, the solid middle line indicates the median value, and lower and upper box edges indicate the interquartile range. Outliers are either less than 3/2 times the lower quartile or greater than 3/2 times the upper quartile. Strong evidence of differing medians is indicated by nonoverlapping notches.

the WDSS-II display during the simulation (Fig. 2a) indicates that the forecaster fixated nearly equally on the northern and southern storms in the reflectivity AOI, whereas he fixated more on the southern storm in the velocity AOI (Figs. 2a and 2b). In Fig. 2c, we see small pockets of longer absolute fixation duration focused on the two storms of interest; however, these pockets are more evident in the velocity AOI. These pockets of longer absolute fixation duration indicate periods of data interrogation focused on specific radar signatures and that the longest fixation duration

(Fig. 4). The interpretation of these trends is aided by computing fixation counts at 5-min intervals, resulting in 8 periods (with the final period being 4 min). These trends are of interest because they indicate variations in the forecaster's cognitive activity. While the forecaster fixated most frequently within the reflectivity AOI (Fig. 4a), the peak fixation count occurred during the fourth period (Fig. 4a). In contrast, the peak fixation count in the velocity AOI occurred in the seventh period and exceeded the corresponding reflectivity AOI fixation count. While in most periods the durations of 5-min-velocity AOI fixations were longest, a minimum in velocity AOI fixation duration occurred in period 4 when fixation duration and fixation counts in the reflectivity AOI were longer and higher (Fig. 4). Like fixation counts in the control AOI, the associated fixation durations were intermittent and tended to be shorter than those in the other 2 AOIs (Fig. 4).

was on signatures within the velocity AOI. Differences in fixation measures between the reflectivity and velocity AOIs suggest that the forecaster used reflectivity data to interrogate both storms and maintain situational awareness of weather within the entire sector, whereas his interrogation of the velocity data was more directed and focused on regions of storms that were of greatest interest.

FIXATION TRENDS.

Trends in the forecaster's fixation counts and mean fixation durations were seen in the 39-min simulation as the weather scenario unfolded

the durations of 5-min-velocity AOI fixations were longest, a minimum in velocity AOI fixation duration occurred in period 4 when fixation duration and fixation counts in the reflectivity AOI were longer and higher (Fig. 4). Like fixation counts in the control AOI, the associated fixation durations were intermittent and tended to be shorter than those in the other 2 AOIs (Fig. 4).

To provide context on how these trends in cognitive activity related to different stages of the forecaster's warning decision process, we created a timeline that summarizes

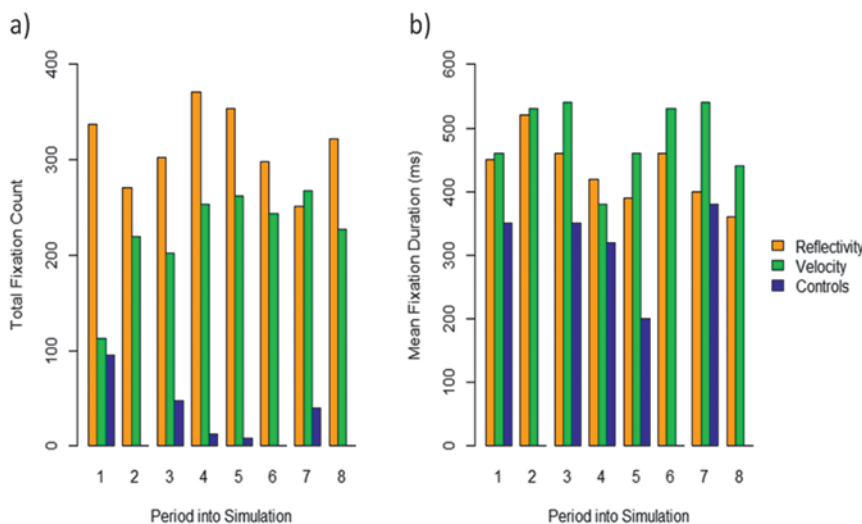


FIG. 4. (a) Total fixation counts and (b) mean fixation durations within the reflectivity (orange), velocity (green), and controls (blue) AOIs per period.

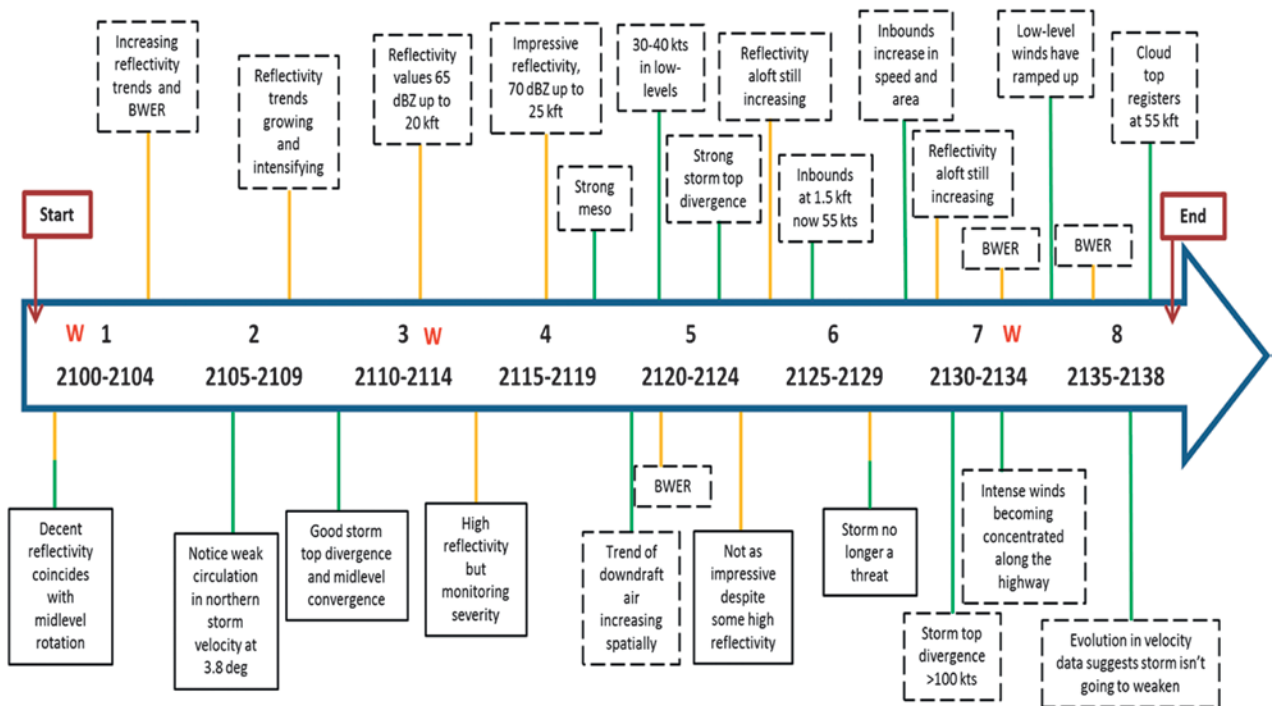


FIG. 5. A timeline of key observations made in the reflectivity (orange) and velocity (green) AOIs with respect to the northern storm (solid box) and southern storm (dashed box). Time period is provided in the arrow for each period (top row) with corresponding case time (UTC) (bottom row). The timings of decisions to issue a warning are indicated by a red “w.”

the forecaster’s retrospective recall during each period (Fig. 5). The initial high number of reflectivity AOI fixation counts in period 1 resulted from using these data to assess storm intensity. After monitoring trends in the height and intensity of the northern and southern storms’ reflectivity cores, the forecaster’s decision to issue a severe warning on the northern storm coincided with the highest peak in the controls AOI fixation count and relatively long fixation durations (Fig. 4a). Similarly, the two other peaks in the controls AOI fixation counts and durations (Fig. 4b) coincided with the issuance of severe weather warnings (Fig. 5). The increasing trend in reflectivity AOI fixation count from a relative minimum in period 2 to its highest peak in period 4 corresponded with the forecaster’s observations of the intensifying southern storm, which he warned on by period 3, and by period 4 he interpreted as being “pretty impressive” with reflectivity values of 70 dBZ up to 25 kft. His focus on reflectivity data also increased because the intensity of the northern storm was diminishing rather than increasing as he had anticipated.

As the southern storm evolved, the downburst potential became apparent to the forecaster, and a

change in his cognitive process was noticeable in both his fixation trends and retrospective recall. During periods 5 through 7, the forecaster’s fixation count in the reflectivity AOI decreased and mean fixation duration in the velocity AOI increased (Figs. 4a and 4b). Concurrently, he began to observe more interesting signatures in the velocity data (Fig. 5). In period 5, he saw a spatial increase in “downdraft air” in the southern storm as well as the presence of “strong cloud-top divergence” (Fig. 5). Although low-level radial winds in the southern storm were only 30–40 kts, the forecaster thought it was “only a matter of time before it really [got] going.” His expectation was confirmed in period 7 when he saw “intense winds becoming concentrated along the highway.” It was also this period that marked the only time that the forecaster’s fixation count in the velocity AOI was higher than in the reflectivity AOI, and the mean fixation duration in the velocity AOI was at a maximum. Following from his observation in the velocity data, he decided to issue a second warning on the southern storm, which corresponds with the third peak in fixation count for the controls AOI (Fig. 4a).

FUTURE APPLICATIONS. The short study presented in this article demonstrates how a forecaster's eye gaze data can be used to understand in greater detail where a forecaster's attention is pointed to and how their attention changed with time. In this instance, we found that the forecaster's fixations changed as a function of the stimulus. We were able to capture his different styles of interrogation of reflectivity and velocity data, and understand how the changing weather scenario impacted the counts and durations of his fixations. Important to our interpretation of trends observed in the fixation measures was the retrospective recall. Together, the eye gaze data and retrospective recall quantified and contextualized the forecaster's cognitive processes, providing a full picture of what, how, and why he was looking at certain points on the screen. The importance of collecting qualitative data to answering the "why" question remains.

The "what" and "how" questions associated with forecasters' decision processes can be answered with more exactness and certainty through eye tracking. Using eye tracking to obtain this more informed knowledge about forecaster decision processes may be useful in a variety of applications within operational meteorology. This informed knowledge will become especially important as efforts to become a Weather Ready Nation continue. For example, Forecasting a Continuum of Environmental Threats (FACETS) is a concept designed to reinvent the watch and warning paradigm from a traditionally deterministic system to one that provides a continuum of probabilistic hazard information (Rothfus et al. 2014). This change in the watch and warning paradigm requires the development and testing of new tools that will meet forecaster needs (Karstens et al. 2015). The widespread application of eye-tracking methods in usability studies (Jacob and Karn 2003) suggests that eye tracking will be useful for learning about forecaster-computer interactions and for successfully designing suitable tools.

Eye tracking may also help determine differences in experienced and expert forecasters' data interrogation strategies and cognitive processes compared to those of the less-experienced forecaster. Understanding these differences would help in the design of effective training for intern and journeymen forecasters. Furthermore, using eye tracking to develop a deeper understanding of forecasters' cognitive processes would be helpful in determining whether new types of data and products support or hinder their warn-

ing decision processes. For example, the impact of higher-temporal resolution radar data on forecasters' warning decision processes has been studied in the Hazardous Weather Testbed (e.g., Heinselman et al. 2015 and Bowden et al. 2015). Recently, eye-tracking was used in the 2015 Phased Array Radar Innovative Sensing Experiment to understand better what these impacts are on forecasters' cognitive processes and their related warning decisions. We expect that collecting forecasters' eye gaze data in addition to their retrospective recalls will better inform us on the specifics of how rapidly updating radar data affect their data interrogation strategies. For example, we will be able to compare trends in fixation measures between forecasters using radar data of differing temporal resolution, analyze their visual scanning patterns, and develop a more complete picture of their decision processes from start to finish. Finally, introducing eye-tracking research methods to operational meteorology studies provides an opportunity for mutual interdisciplinary knowledge growth between the human factor and meteorology research fields, which can only push the boundaries of our current knowledge.

ACKNOWLEDGMENTS. We thank the NWS forecaster who participated in this study and the many NSSL IT experts who helped us overcome technical challenges. Our thanks are also extended to Kurt Hondl, Tanya Schoor, and three anonymous reviewers for providing helpful feedback during the writing of this article. Funding was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement NA11OAR4320072, U.S. Department of Commerce.

FOR FURTHER READING

- Bowden, K. A., P. L. Heinselman, D. M. Kingfield, and R. P. Thomas, 2015: Impacts of phased array radar data on forecaster performance during severe hail and wind events. *Wea. Forecasting*, **30**, 389–404, doi:10.1175/WAF-D-14-00101.1.
- Calhoun, K. M., T. M. Smith, D. M. Kingfield, J. Gao, and D. J. Stensrud, 2014: Forecaster use and evaluation of real-time 3DVAR analyses during severe thunderstorm and tornado warning operations in the Hazardous Weather Testbed. *Wea. Forecasting*, **29**, 601–613, doi:10.1175/WAF-D-13-00107.1.
- Daipha, P., 2015: From bricolage to collage: The making of decisions at a weather forecast office. *Sociol. Forum*, **30**, 787–808, doi:10.1111/sof.12192.

- Drost, R., J. Trobec, C. Steffke, and J. Libarkin, 2015: Eye tracking: Evaluating the impact of gesturing during televised weather forecasts. *Bull. Amer. Meteor. Soc.*, **96**, 387–392, doi:10.1175/BAMS-D-13-00217.1.
- Heinselman, P. L., and S. M. Torres, 2011: High-temporal-resolution capabilities of the National Weather Radar Testbed phased-array radar. *J. Appl. Meteor. Climatol.*, **50**, 579–593, doi:10.1175/2010JAMC2588.1.
- , D. LaDue, D. M. Kingfield, and R. Hoffman, 2015: Tornado warning decisions using phased array radar data. *Wea. Forecasting*, **30**, 57–78, doi:10.1175/WAF-D-14-00042.1.
- Jacob, R. J., and K. S. Karn, 2003: Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, J. Hyönä, R. Radach, and H. Deubel, Eds., Elsevier Science, 573–605.
- Just, M. A., and P. A. Carpenter, 1976: Eye fixations and cognitive processes. *Cognit. Psychol.*, **8**, 441–480, doi:10.1016/0010-0285(76)90015-3.
- Kang, Z., and S. J. Landry, 2014: Using scanpaths as a learning method for a conflict detection task of multiple target tracking. *Hum. Factors*, **56**, 1150–1162, doi:10.1177/0018720814523066.
- 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, doi:10.1175/WAF-D-14-00163.1.
- LaDue, D. S., P. L. Heinselman, and J. F. Newman, 2010: Strengths and limitations of current radar systems for two stakeholder groups in the southern plains. *Bull. Amer. Meteor. Soc.*, **91**, 899–910, doi:10.1175/2009BAMS2830.1.
- Lindell, M. K., and H. Brooks, 2013: Workshop on Weather Ready Nation: Science imperatives for severe thunderstorm research. *Bull. Amer. Meteor. Soc.*, **94**, ES171–ES174, doi:10.1175/BAMS-D-12-00238.1.
- NOAA, 2015: Weather Ready Nation. Accessed 25 November 2015. [Available online at www.nws.noaa.gov/com/weatherreadynation/#.Vk5N5vIVhBd.]
- Olsen, M., 2012: The Tobii I-VT fixation filter: Algorithm description. Accessed 12 November 2015. [Available online at https://stemedhub.org/resources/2173/download/Tobii_WhitePaper_TobiiIVTFixationFilter.pdf.]
- Poole, A., and L. J. Ball, 2006: Eye tracking in HCI and usability research: Current status and future prospects. *Encyclopedia of Human-Computer Interaction*, C. Ghaoui, Ed., Idea Group Inc., 211–219.
- Quoetone, E., J. Boettcher, and C. Spannagle, 2009: How did that happen? A look at factors that go into forecaster warning decisions. Preprints, *34th National Weather Association Annual Meeting*, Norfolk, VA. [Available online at www.nwas.org/meetings/nwa2009/.]
- Rothfus, L. P., C. Karstens, and D. Hilderband, 2014: Next-generation severe weather forecasting and communication. *Eos, Trans. Amer. Geophys. Union*, **95**, 325–326, doi:10.1002/2014EO360001.
- Sherman-Morris, K., K. B. Antonelli, and C. C. Williams, 2015: Measuring the effectiveness of the graphical communication of hurricane storm surge threat. *Wea. Climate Soc.*, **7**, 69–82, doi:10.1175/WCAS-D-13-00073.1.
- Wood, G., K. M. Knapp, B. Rock, C. Cousens, C. Roobottom, and M. R. Wilson, 2013: Visual expertise in detecting and diagnosing skeletal fractures. *Skeletal Radiol.*, **42**, 165–172, doi:10.1007/s00256-012-1503-5.
- Zrnić, D. S., and Coauthors, 2007: Agile beam phased array radar for weather observations. *Bull. Amer. Meteor. Soc.*, **88**, 1753–1766, doi:10.1175/BAMS-88-11-1753.

AMS OFFERS INTRODUCTORY UNDERGRADUATE COURSE PACKAGES TO FIT YOUR INSTRUCTIONAL NEEDS



"The "real time" content and supporting curriculum are powerful teaching tools, especially with the majority of my students majoring in non-STEM and non-geoscience disciplines. Many of them take the course as a "general elective" with little expectation of becoming interested. However, many of them have reported being significantly more aware of how much atmospheric science impacts their everyday lives."

– Dr. David A. Padgett, Tennessee State University



Weather
STUDIES

Photo: Anupa Asokan



Ocean
STUDIES

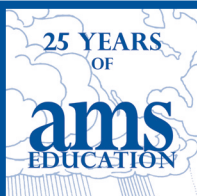
Photo: Anupa Asokan



Climate
STUDIES

Photo: Jean-Baptiste Dodane

- ✓ **REAL-WORLD LEARNING EXPERIENCES FROM THE COMMUNITY YOU TRUST, AMERICAN METEOROLOGICAL SOCIETY**
- ✓ AMS members contribute to the editing and creation of course material, including dynamic eTextbooks
- ✓ eTextbooks, eInvestigation Manual, Real-Time Course Portals to supplement your curriculum
- ✓ You can now read our books and manuals online on any device without downloading!



REQUEST A FREE EXAM COPY TODAY!
ametsoc.org/UndergradFaculty