

Maintaining the Role of Humans in the Forecast Process

Analyzing the Psyche of Expert Forecasters

BY NEIL A. STUART, DAVID M. SCHULTZ, AND GARY KLEIN

At the Second Forum on the Future Role of the Human in the Forecast Process, held on 2–3 August 2005 at the AMS Weather and Forecasting Conference in Washington, D.C., an international group from all sectors of our community gathered to address a number of issues, including the cognitive and psychological aspects of expert weather forecasting. Like the first forum (discussed by Stuart et al. in a 2006 *Bulletin* article, pp. 1497–1502), the second forum provided opportunities for a broad cross-section of the profession to discuss timely issues in forecasting. This time, we hoped to find insight into how people might maintain a crucial role in forecasting by examining the psychological aspects of training forecasters, as well as examining the forecast process itself.

BRIDGING THE GAP BETWEEN THEORY AND APPLICATION. One illustration of the role of psychology in forecasting can be found in studies of the initial education of meteorologists. Between 1981 and 1995, the number of B.S. degrees awarded in physical sciences decreased 20%, despite growth in the overall number of undergraduates. Undergraduate

atmospheric science programs typically experience a 30%–50% attrition rate from the time freshman enter college to graduation (P. Roebber and F. Carr 2005, personal communication). The number of applications to graduate school in atmospheric science has decreased, especially from domestic students. In addition, faculty are often dissatisfied with student learning and motivation. In a 2005 *Bulletin* article (pp. 507–517), Roebber suggested that these statistics might indicate a gap between student and faculty expectations.

In an analogy to numerical weather prediction, Roebber asserted that limitations to instructional success were related to differences between the initial states of students and faculty and imperfect teaching models. Educational research suggests that students are typically goal-seeking learners, requiring concrete examples, case studies, and applications of theory. In contrast, professors are knowledge-seeking learners, preferring to understand the theory before applying it to experiences. This conflict between student and teacher learning styles leads to suboptimal learning. Roebber surveyed undergraduate students, graduate and postgraduate students, and faculty at University of Wisconsin—Milwaukee to confirm this educational divide between learning styles.

A similar disconnect occurs between researchers and forecasters. While researchers prefer to understand theoretical concepts (knowledge seekers), forecasters prefer concrete examples from the real world (goal seekers). In order to optimize improvements in weather forecasting, those involved in forecaster education need to adapt to this divide so that goal seekers can learn most effectively. Roebber suggested that professors present theoretical concepts with examples from the real world, providing ample time for active student participation, encouraging students to work in teams on assignments, and connecting current course material to past material, other courses, or everyday life.

AFFILIATIONS: STUART—NOAA/National Weather Service, Wakefield, Virginia (*Current affiliation:* National Weather Service, Albany, New York); SCHULTZ—Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/National Severe Storms Laboratory, Norman, Oklahoma (*Current affiliation:* Division of Atmospheric Sciences, Department of Physical Sciences, University of Helsinki, and Finnish Meteorological Institute, Helsinki, Finland); KLEIN—Klein Associates Division of Applied Research Associates, Fairborn, Ohio

CORRESPONDING AUTHOR: Neil A. Stuart, National Weather Service, 251 Fuller Road, Suite B300, Albany, NY 12203
E-mail: Neil.Stuart@noaa.gov

DOI:10.1175/BAMS-88-12-1893

©2007 American Meteorological Society

HEURISTICS IN DECISION MAKING. Understanding the psychology of decision making would help forecasters improve beyond their meteorological knowledge and experience.

Specifically, weather forecasting involves many difficult decisions. Forecasters face incomplete or inaccurate diagnoses of the atmosphere, multiple sources of imperfect computerized guidance, and difficulty in using the immense volumes of information available. In severe weather, individual forecasters have different thresholds for issuing warnings. Most forecasters believe the penalty for not forecasting severe weather is greater than the penalty for issuing a watch/warning that proves false. This increases the likelihood of false warnings, and although recent research in a 2007 *Weather and Forecasting* article by Barnes et al. (pp. 1140–1147) suggests a significant segment of the user community does not become desensitized by false alarms, eventually some segments of the user community may not take action when necessary.

Forecasters are faced with so much information from observations and from multiple numerical model forecasts that in some circumstances an analytical decision-making process is limited by time constraints. Thus, heuristic approaches are virtually mandatory for weather forecasters. In a 2004 *Weather and Forecasting* article (pp. 1115–1126), Doswell addressed the heuristics involved in forecasting and decision making by humans under time constraints and uncertainty. Heuristics—intuitive methods—allow forecasters to absorb diverse and even incomplete information to arrive at forecast decisions. Although heuristics aren't foolproof and are usually based on individual experience and perceptions, they allow people to improve upon purely objective forecasts. Purely analytical methods, as discussed by Doswell, produce forecasts that tend to cluster narrowly around the verifying data, but occasionally can produce very bad forecasts. Purely intuitive forecasts, on the other hand, typically have a relatively broad distribution around the verifying data that can produce different kinds of errors. Most forecasters use a mix of intuitive and analytic methods. These forecasts also can be substantially better than those produced by purely objective methods or purely intuitive methods, but can occasionally result in poor forecasts.

Many forecasters, especially experienced forecasters, use conceptual models and pattern recognition in heuristic ways. That is, although these methods may be based on analytical or empirical data and experi-

ence, people apply these intuitively when making forecast decisions. Broadly speaking, these are based on mental models derived from past experience, along with empirical evaluation of past forecasts. Much remains to be learned about precisely how consistently good forecasters use the information they have, but it is clear that the best human forecasters can consistently outperform objective methods (such as model output statistics, MOS), helping to ensure that people will have a role in the forecast process.

All forecasters have limited capabilities for analytical reasoning and correspondence-based judgment, in part owing to a lack of adequate training and education, but also due to time constraints. If forecasters continue to make only minimal improvements over MOS, humans may eventually disappear from the forecast process.

EXPERIENCE, INEXPERIENCE, AND FLEX-

IBILITY. Klein Associates has studied the cognitive and psychological aspects of professions that require quick and effective decisions, such as the military, firefighting, and nursing. Recently, Klein Associates has studied weather forecasters. Based on these studies, forecasters can be divided into five categories.

- a) Intuitive scientists—innovative, creative, and decisive
- b) Rule-based scientists—minimal creativity, reliance on guidance, less reliance on intuition
- c) Procedure-based forecasters—confined to routines, little flexibility in unusual situations
- d) Procedure-based mechanics—concerned only about product formats and deadlines
- e) Disengaged—little interest in the job

Klein Associates' research showed that in a very small sample of U.S. military meteorologists, nearly 40% could be categorized as procedure-based mechanics or disengaged, treating their jobs mainly as a source of income. The disengaged and procedure-based mechanics seem to exist in all sectors of the profession, though no formal research has quantified this.

Klein Associates showed distinct differences between inexperienced and experienced (expert) forecasters. Inexperienced forecasters rely too much on computer models, have a narrow perspective, use a fixed set of procedures, and are reactive, usually resorting to "chasing the observations." Conversely, expert forecasters routinely identify the significant problem of the day and use a more global perspective

when initially diagnosing the state of the atmosphere. An expert uses tools and procedures flexibly. They use mental or conceptual models to interpret the weather scenario.

Expert forecasters employ a recognition-primed decision model (Klein 1998), which combines both analysis and intuition in their decision-making process. This model overcomes limitations to effective decision making through direct analysis (e.g., time constraints, too much data, other distractions). For example, a weather scenario exhibits certain cues that can trigger a mental simulation of a past weather event. Both analysis and intuition are needed for effective decision making.

THE COGNITIVE DIMENSION OF TECHNOLOGY. Overreliance on technology (e.g., guidance from numerical weather prediction models, new computer hardware and software, automated forecast production) to reduce uncertainty can often contribute to less effective decision making. Klein's results show that experienced and inexperienced forecasters will respond differently to the introduction of information technology (IT). IT raises the level of the inexperienced forecasters by providing a convenient crutch for saving time. On the other hand, IT can inhibit experienced forecasters, making them less adaptive and more passive, pigeonholing or disconnecting them from the way they prefer to analyze the data. This disrupts the mental simulation and recognition-primed decision modeling process important to intuitive-scientist forecasters. Sometimes, however, experienced forecasters flourish with IT tools, analyzing weather scenarios in new and creative ways to enhance their understanding.

Supporting these claims was Project Phoenix, a research effort by Purcell, McCarthy, and others at Environment Canada's Prairie Storm Prediction Center in Winnipeg, Manitoba. Forecasts from teams of people were verified against an automated forecasting system (see McCarthy et al. 2007 for more details about Project Phoenix). Each team's forecasts made significant improvement over the automated system up to 24 h from the initial time, with dramatic improvements within 12 h. The team forecasts produced fewer errors than the automated forecasts, likely because of the perceived competition between the teams and the automated forecasts. Problems with the automated forecasts included errors in resolving initial conditions that grew exponentially with time. (Data that have not been checked for accuracy or

representativeness can be included in forecast model initializations, producing errors that grow with time.) These findings supported the concept that availability of short-term model data of any kind may be detrimental to accuracy if the model data are erroneous.

Furthermore, increasing the volume and sophistication of IT puts more emphasis on improving computer guidance and automating procedures for producing forecast products. Information technology can contribute to forecaster apathy, if a forecaster is resigned to the fact that value cannot be added to computer guidance. Forecasters not motivated to analyze data and computer guidance, to add value to the computer guidance, cannot learn from the past and build expertise. Apathetic forecasters can all too easily fall into the 40% of the workforce that are procedure-based mechanics or disengaged and can be the least effective decision makers. Meanwhile, experienced forecasters are often motivated to optimize new methods of data analysis, computer guidance, and product development. Additionally, they find ways to continue to gain expertise as their role in the forecast and decision-making process evolves.

CONCLUSION. These psychological considerations emphasize the importance of proper forecaster education in building quality experience so that humans remain important in the weather forecasting process. Such education includes formalized training on new concepts, training simulations on past weather events, or training hardware and software in the forecast office. Understanding of heuristics and cognitive considerations could guide people designing and implementing educational tools to most effectively build experience in forecast staff. Educators would benefit from results of research on what distinguishes intuitive-scientist forecasters from disengaged forecasters, which in turn could result in the development of training materials to effectively guide forecasters in their pursuit of experience and expertise.

As discussed in Doswell's article, collaboration is needed among meteorologists, cognitive psychologists, and others studying decision making to study the heuristics involved in decision making under time constraints and uncertainty, as well as evolving information technology. Studying the cognitive styles of judgment and decision making of expert forecasters versus mediocre/poor forecasters may facilitate improvements and will help illuminate possible future roles for humans in the forecast process.

Finally, cognitive styles of forecasters go beyond the length of forecasting experience. Steps should be taken to quantify the quality—rather than the length—of experience to identify the strongest forecasters. Length of service does not always equal quality of experience if a forecaster does not learn from past events and improve their forecasting skills. Ultimately, methods to motivate forecasters to increase their expertise and improve their decision-making abilities will help forecasters to continue to add value to computer-generated output, and likely ensure the future of people in the forecast process.

ACKNOWLEDGMENTS. The authors would like to thank Mary Cairns and the AMS Weather and Forecasting Committee for their support in planning this event. Additionally, we would like to thank the staff of NWS Wakefield, Virginia; the National Severe Storms Laboratory; and Klein Associates for providing the facilitators and presenters for this event. Comments by Paul Roebber of the University of Wisconsin—Milwaukee during the preparation of Schultz’s presentation were quite valuable. Thanks also to Chuck Doswell of the Cooperative Institute for Mesoscale Meteorological Studies, and Dan Reilly and John Billet of NWS Wakefield, Virginia, for their assistance in revising

REACHING THE GOALS OUTLINED IN THE FIRST FORUM

While one session at the Second Forum on the Future Role of the Human in the Forecast Process addressed the psychological aspects of effective forecasting (discussed in the accompanying essay), the other sessions defined steps needed to reach the goals of the first forum (discussed by Stuart et al. in a 2006 *Bulletin* article, pp. 1497–1502). Four major themes identified at the previous forum were:

- improved communication within the profession and to users;
- improved task delegation and time management in forecast production;
- increased understanding and usage of probabilistic forecasts; and
- our ability to improve upon numerical weather prediction (NWP) model guidance.

The discussion brought forth interesting follow-up ideas about credible and meaningful ways of expressing uncertainty and educating people about it.

Whether in words or graphics, the uncertainty or specificity of forecasts must be described. Improved communication could mean expressing probabilities of when the greatest impact of specific weather hazards is expected, or the ranges of possible extremes. Not all users have the same level of sophistication. For example, leaders of the annual national Boy Scout Jamboree, a summertime gathering of thousands of people, need to be apprised of the heat index and how it will likely affect their plans, as well as the likelihood and timing of thunderstorms and the threat of lightning. We need to educate emergency managers, civic leaders, and others who lack specialized training by avoiding scientific jargon and describing weather impacts.

We need to be “ambassadors of science” and show that probabilistic forecasts can offer more information than deterministic forecasts. In particular, probabilistic forecasts can be used in cost–loss models; however such models might make forecasters with more certainty seem more accurate. Furthermore, forecasts without objectively verified accuracy

can lead to misleading results. The more our customers know, the more effective our message.

Through partnerships between academia, NWS, the media, emergency managers, and schools, we can inform and educate our customers; for instance, the media can experiment with and describe new products. Private companies can make a similar investment to educate their users.

Forecast organizations should strive to understand customers’ needs and ask them directly, “Are you satisfied with my (our) products?” If customers seek inappropriate information, we must determine how to best meet their needs. A “customer satisfaction index” could be developed through user feedback, perhaps broken down into relevant categories. By polling our users every few years, we can gauge how well we are doing. Similarly, our product development should involve users.

The group was evenly split about the importance of displaying more probabilistic information in forecasts. Media and private forecasters are often pressured to present a sense of accuracy and certainty. Many believe a single temperature forecast value is more understandable than a range of values. The consensus was that some probabilistic products work well, such as for hurricane strike and severe weather and flood potential. We could build upon public acceptance of these forecasts; the next step could be to apply strike probabilities to heavy snow and ice forecasts. There was no agreement on how to approach probabilistic temperature forecasts.

Despite studies to the contrary (such as the 1980 *Bulletin* article by Murphy et al., pp. 695–701), some forecasters feel few people know what a “POP (probability of precipitation)” forecast is, including many forecasters! Training forecasters to use probabilistic forecasts is just as important as educating users, so forecasters can convey information with a coherent voice. On Web-based AMS forums, forecasters could share ideas and experience and feel “safe” to ask questions. Chat rooms, blogs, and/or email distribution lists for forecasters could be developed.

the manuscript. Funding for Schultz was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA17RJ1227, U.S. Department of Commerce.

Additional thanks go to Ed Johnson, director of the Strategic Planning and Policy Office at NWS Headquarters in Silver Spring, Maryland, for his participation in the forum.

Finally, the authors would like to offer a special acknowledgment to forum facilitator and coauthor James Moore of St. Louis University, St. Louis, Missouri, who passed away in July 2006. Participation in this forum was one of his final appearances at a special session at an AMS conference. His

dedication, extensive knowledge, and enthusiasm about all aspects of atmospheric research and operational meteorology, and his sense of humor, are greatly missed.

The views expressed are those of the authors and do not necessarily represent those of the National Weather Service.

FOR FURTHER READING

Andra, D. L. Jr., E. M. Quetone, and W. F. Bunting, 2002: Warning decision making: The relative roles of conceptual models, technology, strategy, and forecaster expertise on 3 May 1999. *Wea. Forecasting*, 17, 559–566.

Communications between the forecasters and researchers should improve to better address the needs of the forecaster and quickly infuse the latest scientific ideas and products into forecast operations.

Training about ensemble data is also important. The evolution of public weather information depends on our ability to leverage information from ensembles. (Some participants focused on ensuring that ensembles adequately represent the range of forecast possibilities. Certain fields, such as temperatures and precipitation, should be calibrated, to be optimized with time. These calibrated products eventually should be the basis of the forecast process.) While about 50% of the participants agreed that the overuse of deterministic NWP guidance is a problem, some participants viewed ensembles as a fad. The consensus was that easy-to-use ensemble displays of uncertainty and probabilities would facilitate the increased use of these products. Documentation on how to use these products will be necessary, and in turn could be a reference in the development of more probabilistic forecast products. Most forecasters already use an ensemble methodology by comparing deterministic models and using model output statistics. Educators can build upon these practices to demonstrate the need for more probabilistic products.

After a broad education at the undergraduate level, forecasters need frequent and improved training on the latest scientific ideas and products. Modules and teletraining should be short and designed in consultation with experienced forecasters. (Similarly, the operational community should be consulted before new forecast methods and products are developed. Too often, decisions are made by people with little or no operational experience.)

Frequent regional workshops and conferences are needed to share experience and demonstrate ideas and concepts. Forecasters need meaningful verification to track their own progress and address deficiencies. Finally, we need seasonal training emphasizing regional weather.

Time management appears to be a universal problem of staffing, information flow, and forecaster knowledge and

experience. Such problems may be alleviated by emphasizing the day's greatest forecast problem (e.g., foregoing extreme detail in a day-6 forecast to focus on a day-2 approaching winter storm). A forecaster determines what products are best suited for a particular situation and how to distribute workload.

A 2006 *Weather and Forecasting* article by Baars and Mass (pp. 1034–1047) shows, in general, that consensus from multiple sources of model output statistics (MOS) or the mean of ensemble model guidance provides the most accurate temperatures and precipitation forecasts beyond a subjective short range. Forum participants noted, however, that correction of model or member bias as well as intimate knowledge of local effects could, for spot deterministic forecasts, outperform the consensus.

Forecasters must be able to identify when anomalous weather necessitates deviation from NWP model output. They can add value to forecasts for the next few hours, most notably in high-impact, low-frequency weather such as severe thunderstorms. In the 3- to 8-day range, for benign weather, NWP is difficult to improve upon. In benign situations, rather than spending time adding value to NWP model output, forecasters should shift toward communication and providing confidence levels for forecasts of hazardous periods.

Adding value to guidance only for hazardous weather, and emphasizing improved communication to users, would be a fundamental shift of operations in both governmental and commercial forecasting. Adding value to NWP model guidance output is one way to ensure the future need for humans in forecasting. Similarly, as long as users need effective communication of uncertainty and levels of confidence, humans are vital to the forecast process.

—NEIL A. STUART, RICHARD GRUMM (NWS), JAMES MOORE (SAINT LOUIS UNIVERSITY), ALBERT E. PIETRZYCHA (NWS), KENNETH REEVES (ACCUWEATHER), ELLIOT ABRAMS (ACCUWEATHER), AND CLIFFORD F. MASS (UNIVERSITY OF WASHINGTON)

- Baars, J. A., and C. F. Mass, 2006: Performance of National Weather Service forecasts compared to operational, consensus, and weighted model output statistics. *Wea. Forecasting*, **20**, 1034–1047.
- Barnes, L. R., E. Grunfest, M. H. Hayden, D. M. Schultz, and C. Benight, 2007: False alarms and close calls: A conceptual model of warning accuracy. *Wea. Forecasting*, **22**, 1140–1147.
- Doswell, C. A. III, 2004: Weather forecasting by humans: Heuristics and decision making. *Wea. Forecasting*, **19**, 1115–1126.
- Hahn, B. B., E. Rall, and D. W. Klinger, 2002: Cognitive task analysis of the warning forecaster task. Final Report for National Weather Service, Office of Climate, Water, and Weather Service, Order No. RA1330-02-SE-0280, 28 pp. [Available from Klein Associates, 1750 Commerce Center Blvd North, Fairborn, Ohio 45324.]
- Klein, G., 1998: *Sources of Power: How People Make Decisions*. The MIT Press, 330 pp.
- McCarthy, P. J., D. Ball, and W. Purcell, 2007: Project Phoenix: Optimizing the machine–person mix in high-impact weather forecasting. Preprints, *22nd Conference on Weather Analysis and Forecasting*, Amer. Meteor. Soc., Park City, UT.
- Murphy, A. H., S. Lichtenstein, B. Fischhoff, and R. L. Winkler, 1980: Misinterpretation of probability of precipitation forecasts. *Bull. Amer. Meteor. Soc.*, **61**, 695–701.
- National Center for Education Statistics, 1998: The condition of education, 1998. National Center for Education Statistics, U.S. Department of Education, NCES 98-013, GPO 065-000-01196-4, 377 pp. [Available online at <http://nces.ed.gov/pubs98/98013.pdf>.]
- Pliske, R. M., D. Klinger, R. Hutton, B. Crandall, B. Knight, and G. Klein, 1997: Understanding skilled weather forecasting: Implications for training and the design of forecasting tools. Tech. Rep. No. AL/HR-CR-1997-0003 for the Air Force Material Command, Armstrong Laboratory, Human Resources 16 Directorate, Brooks AFB, TX, 122 pp. [Available from Klein Associates, 1750 Commerce Center Blvd North, Fairborn, Ohio 45324.]
- , B. Crandall, and G. Klein, 2004: Competence in weather forecasting. *Psychological Investigations of Competence in Decision Making*, K. Smith, J. Shanteau, and P. Johnson, Eds., Cambridge University Press, 40–68.
- Roebber, P. J., 2005: Bridging the gap between theory and applications: An inquiry into atmospheric science teaching. *Bull. Amer. Meteor. Soc.*, **86**, 507–517.
- Stuart, N. A., and Coauthors, 2006: The future role of the human in an increasingly automated forecast process. *Bull. Amer. Meteor. Soc.*, **87**, 1497–1502.
- Vali, G., R. Anthes, D. Thomson, D. Houghton, J. Fellows, and S. Friberg, 2002: Wanted: More Ph.D.s—Graduate enrollments in the atmospheric sciences. *Bull. Amer. Meteor. Soc.*, **83**, 63–71.
- Verret, R., G. Babin, D. Vigneux, J. Marcoux, J. Boulais, R. Parent, S. Payer, and F. Petrucci, 1997: Scribe: An interactive system for composition of meteorological forecasts. Preprints, *11th International Conf. on Interactive Processing Systems for Meteorology, Oceanography, and Hydrology*, Amer. Meteor. Soc., Dallas, TX, 56–61.