Chapter 15

Computer-Aided Visualization in Meteorology

J. Gregory Trafton
Naval Research Laboratory, Washington, DC

Robert R. Hoffman
Institute for Human and Machine Cognition, Pensacola, FL

Our topic in this chapter is not so much what happens when experts have to work "out of context," but how cognitive engineering might help weather forecasters, in particular, remain within familiar decision-making spaces by improving on their display technology. Ballas (chap. 13, this volume) makes it clear how weather forecasting, as a workplace, is a constantly moving target by virtue of continual change in displays and data types. Since weather maps were invented around 1816, meteorological data visualization has gone through many dramatic changes (Monmonier, 1999). Weather maps are now displayed and manipulated by computer, even though hand chart work is still generally regarded as a critical activity in the forecaster's trade (see Hoffman & Markman, 2001). Most weather forecasters get data, charts, and satellite images from Internet sources including the World Wide Web. In this chapter, we discuss some of what we know about how weather forecasters use information technology to display and support the interpretation of complex meteorological visualizations. Based on notions of human-centered computing (HCC), we offer some suggestions on how to improve the visualizations and tools.

Norman (1993) offered design guidance that would (hopefully) ensure human-centeredness. Ideally, such guidance should be applied throughout the entire design process, but this does not always happen. In the case of weather forecasting, some of the ways of representing data (such as wind-barbs or iso-pressure lines) were standardized long ago. Ingrained traditions in meteorological symbology and display design force one to keep in mind

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE **2007**	2. REPORT TYPE	3. DATES COVERED **00-00-2007 to 00-00-2007**

4. TITLE AND SUBTITLE **Computer-Aided Visualization in Meteorology**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory,Navy Center for Applied Research in Artificial Intelligence (NCARAI),4555 Overlook Avenue SW,Washington,DC,20375**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**22**	

what we call the Will Robinson principle. This principle is named after a character in the television series *Lost in Space* who was continually cautioned by his robot about impending dangers. The idea of the principle is that it is potentially dangerous for the outsider—a researcher who is analyzing a domain to create new technologies—to inject changes in long-standing traditions, no matter how flawed the traditions may seem at first glance.

One of our first experiences that (eventually) spawned this principle involved studies of expertise in the field of remote sensing in general (Hoffman, 1990; Hoffman & Conway, 1990) and a study of expertise at interpreting infrared satellite imagery (Hoffman, 1997; Klein & Hoffman, 1992). Until color was introduced, infrared weather satellite images were portrayed using a gray scale representing temperatures. On first thought, one might think that cold things would be represented as black and dark gray tones, and relatively warmer things as light tones to white. However, the Earth seen in this tonal scale looks like a surreal marble cake, not even much like a planet. Reversing the scale (warm = black, cold = white) suddenly makes the Earth look like the Earth and the clouds look like clouds. One of the tonal scales (or "enhancement curves") that forecasters found valuable (and still use to map temperatures, though color has been added), uses black to dark gray tones for the relatively warmest temperatures (i.e., the land surface), lighter grays for the lower and relatively warm clouds, but then reverts back to dark gray for the relatively cooler midlevel clouds, proceeding up through the lighter gray shades to white, and then yet again back to dark gray and black for the relatively coolest and highest cloud tops. An example is shown in Figure 15.1.

Taken out of context as a tone-to-temperature mapping scale, it is somewhat mysterious to the outsider. But with practice comes a skill of being able to use the repeating ascending tone scales to perceive cloud height and thereby gain an awareness of atmospheric dynamics. (Higher cloud tops are relatively colder and also more massive, thus representing the presence of greater amounts of moisture that can be associated with storms and precipitation at ground levels. Note the storm cell in the lower left of Figure 15.1.) It was only after a wave of cognitive task analysis with expert forecasters that the value of the enhancement curve became apparent.

With this cautionary tale in mind, the question we ask in this chapter is: "Given the current state of weather visualizations, how can we apply psychological and human-centering principles to improve the forecasting process?" In order to answer this question, we briefly describe two of the "design challenges" of HCC (see Endsley & Hoffman, 2002)—the Lewis and Clark Challenge and the Sacagawea Challenge, and provide examples from studies supported by the Office of Naval Research that suggest how we can apply the notions of human centering to improve the visualizations that forecasters use, and the forecasting process itself, to help forecasters stay "in context."

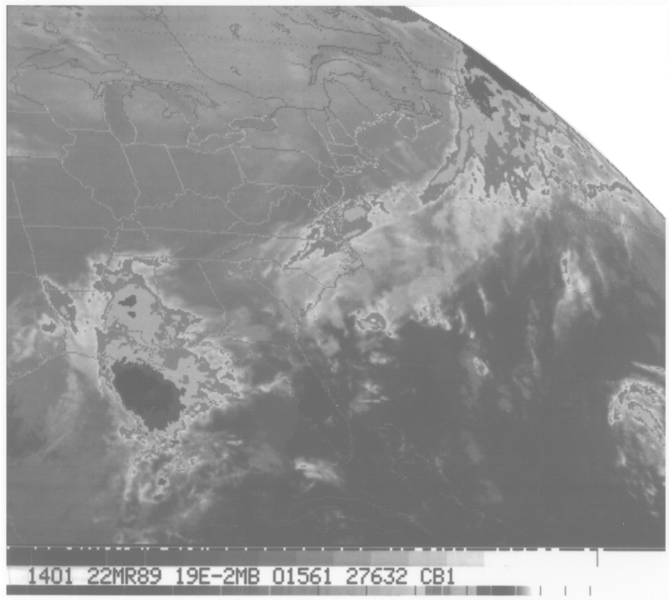

FIG. 15.1. A representative image generated by applying the "MB enhancement curve" to infra-red radiometric data sensed by GOES (Geostationary Operational Environmental Satellite) of the National Oceanographic and Atmospheric Administration. The upper gray scale in the legend at the bottom maps temperatures onto a continuous tonal palette. The bottom gray scale is the enhancement curve. The upper right corner appears cropped but is the apparent temperature of space, coded as white (i.e., cold).

TWO PRINCIPLES OF HUMAN-CENTERED COMPUTING

How should information be displayed on a computer screen? The answer is, of course, "It depends." What we want to focus on is the fact that the easiest way to display information is not necessarily the easiest way for a person to understand the information. Presenting external information in a way that people find meaningful is the key to what Donald Norman (1993) calls "naturalness of representation":

> Perceptual and spatial representations are more natural and therefore to be preferred but only if the mapping between the representation and what it stands for is natural—analogous to the real perceptual and spatial environment . . . the visible, surface representations should conform to the forms that people find comfortable: names, text, drawings, meaningful naturalistic sounds, and perceptually based representations. The problem is that it is easiest to pres-

> ent people with the same representations used by the machines: numbers. This is not the way it ought to be. Sure, let the machines use numbers internally, but present the human operators with information in the format most appropriate to their needs and to the task they must perform. (pp. 72, 226)

It is a strong claim indeed that displays should use representations that match the user's mental representations, and it by no means clear that Norman's (1993) sweeping generalization is the right way to go. The simplest case, the case lying at the heart of Norman's treatment of this issue, is to present information graphically if people understand it visually or in the form of mental imagery, rather than presenting tables showing the numbers that computers so adeptly process. But beyond this simplest case is a large murky zone in which one cannot always be certain how the human mind represents things, or how weather displays should somehow align with mental representations. A case in point might be the example given earlier, of the tonal enhancement curves for satellite imagery. As Endsley and Hoffman (2002) point out, the core notion for which Norman is reaching seems to be that displays need to present meanings, and do so in a way that is directly perceptible and comprehensible.

But ease, directness, and psychological reality are differing meanings of "natural." What we might call immediately interpretable displays might, but need not necessarily, present information in the same way it is represented in the human mind or manifested in the "real world." On the contrary, displays of data from measurements made by individual sensors may be computationally integrated into goal-relevant higher order dynamic invariants or compound variables.

A good example from meteorology is the "skew-T, log p" diagram. This diagram uses a clever trick—measuring elevation or height in terms of pressure. This makes the y-axis rather like a rubber ruler. When a mass of air has relatively high pressure, the ruler is scrunched; and when a mass of air has relatively low pressure, the ruler is stretched. Temperature is the x-axis, and the interpretation of the diagram involves looking for patterns that appear as changes in temperature (skews) as a function of height in the atmosphere as measured in terms of pressure ("geopotential height"). To those who are unfamiliar with the skew-T diagram, its appearance and interpretation are a mystery. To those who are familiar with it, the diagram provides immediately perceptible clues to atmospheric dynamics.

Stimulated by Norman's guidance, Endsley and Hoffman (2002) offered two related principles of HCC. One is the Sacagawea Challenge, named after the guide for the Lewis and Clark Expedition:

> Human-centered computational tools need to support active organization of information, active search for information, active exploration of information,

> reflection on the meaning of information, and evaluation and choice among action sequence alternatives.

The Lewis and Clark Challenge states that:

> The human user of the guidance needs to be shown the guidance in a way that is organized in terms of their major goals. Information needed for each particular goal should be shown in a meaningful form, and should allow the human to directly comprehend the major decisions associated with each goal.

Before we discuss the application of these design challenges to meteorological visualization, we need to present a primer about forecasting. (For a discussion of the human factors in general remote sensing, see Hoffman, 1990; Hoffman & Conway, 1990; Hoffman & Markman, 2001.)

WEATHER-FORECASTING TOOLS AND TECHNOLOGIES

In a nutshell, the forecaster's job is to use weather data (observations, radar, satellite, etc.) and the outputs of the many weather-forecasting computer models to make accurate products, including forecasts, guidance for aviation, flood warnings, and so on. Observational data available to the forecaster include ground-based observations from specialized sensor suites located at airports and forecasting facilities (winds, precipitation, air pressure, etc.). Data are also provided by balloon-borne sensors that are launched from weather-forecasting facilities.

The computer models for weather forecasting rely on both the statistics of climate and physical models of atmospheric dynamics to go from current observational data (which are used to "initialize" the physical models) to guidance about what the atmospheric dynamics will be at a number of scales of both space and future time. The computer models make forecasts only in a limited sense. No one can be a really good weather forecaster by relying on the computer model outputs unless she or he can forecast the weather without using the computer model outputs. The computer models are not infallible. Indeed, they are often "supervised" (that is, tweaked; see Ballas, chap. 13, this volume). The data used for initialization may not be as timely or reliable as the forecaster might like. The different computer models (based on different subsets of physics) often make different weather predictions and all have certain tendencies or biases (e.g., a model may tend to overpredict the depth of low-pressure centers that form over the eastern U.S. coastline after "skipping over" the Appalachian mountains). This has led to the creation of "ensemble forecasts" that integrate the outputs from a number of the individual computer models. On the one hand, this adds to the forecaster's toolkit, but on the other, it adds to the data overload problem.

Satellite images show both past and recent truth and provide the forecaster with the "big picture" of dynamics at a global scale. There are numerous types of satellite data from many platforms including GOES, SEASAT, and others. The image products portray a variety of data types, including visible, infrared, water vapor, and so forth. Radar, especially from the remarkable NEXRAD system, provides the forecaster with a great deal of information, including winds, precipitation, and so on. The NEXRAD is linked into offices of the National Weather Service, where forecasters can request any of a great variety of products from the radar products generator (velocity data, reflectivity, precipitable water, etc.).

New software tools allow the forecaster to combine data types into single displays. For instance, a map showing "significant weather events " (e.g., storms) may have overlaid on it the data from the national lightning-detection network. An image from the GOES satellite may have overlaid on it the data from the NEXRAD radar, and so on.

As one might surmise from this brief presentation, the typical weather-forecasting facility centers around work areas populated by upwards of a dozen workstations displaying various data types. Staff includes forecasters responsible for general forecasting operations but also forecasters responsible for hydrometeorology and for aviation. At any one, time a forecaster might decide to examine any of scores of different displays or data sets (Hoffman, 1991; Trafton, Marshall, Mintz, & Trickett, 2002). Despite the recent introduction of all the new software and hardware systems, the typical weather-forecasting facility still finds need for such things as chart tables, clipboards, and colored markers. It is by no means clear whether, how, and to what degree any of the technologies either supports or handicaps the forecasting process. One might therefore legitimately ask whether the forecasting facilities might benefit from the application of the notions of HCC.

APPLYING THE PRINCIPLES

When a system is designed, how are the types of displays or interfaces chosen? Frequently, decisions are based on what is easiest or most efficient from a programming point of view with little regard to people actually perform the task. The Lewis and Clark Challenge and the Sacagawea Challenge suggest that the goals and mental operations of the human are critical to the success of the computational tools or displays. After determining the human's reasoning, the designer can try to determine how well (or poorly) an interface supports (or prevents) it. The designer can then change or even completely redesign the interface to facilitate both frequent and difficult mental operations, and then engage in an empirical usability evaluation.

A necessary and often difficult part of this task, of course, is revealing and describing the forecaster's reasoning. There are many methods of cognitive task analysis including methods for modeling cognition at the microscale of the reaction times and sequences of individual mental operations (e.g., GOMS). There are several tutorials available for conducting cognitive task analysis (e.g., Crandall, Klein, & Hoffman, 2006; Hoffman & Lintern, 2006; Schraagen, Chipman, & Shalin, 2000; Trafton et al., 2002; Vicente, 2000). There are also ways of revealing the perceptual steps a user goes through (e.g., with an eye tracker) in conducting particular tasks. Forecasting also has to be described at a macrocognitive level of high-level strategies and the drive to comprehend of what's going on in the atmosphere. One thing that makes weather forecasting a challenging domain for HCC analysis is that there is no one single model of how forecasters reason. Reasoning methods and strategies depend on climate, season, experience level, and a host of other factors. (A detailed discussion appears in Hoffman, Trafton, & Roebber, 2006.)

The studies we now discuss examined how experienced meteorologists understand the weather and create a weather forecast. The forecasters used a variety of computer tools including internet sources, meteorology-specific tools, and off-the-shelf tools such as Microsoft PowerPoint. They also communicated with other to make sure their forecast and understanding of the current weather conditions was correct.

How might we apply the Sacagawea Challenge to weather forecasting? Norman's (1993) guidance, which we have already qualified previously, suggests that weather-forecasting displays that show graphics (images, charts) would already be "natural," and those that depict quantitative information would be "not-natural." Here we see how Norman's distinction is of little help, because weather charts are anything but "natural" and because forecasters need to know quantitative values (e.g., wind speeds at a height of 700 millibars) and must have some idea of specific likelihoods, certainly by the time they finish a forecast. For example, a forecaster may predict that there will be a 30% chance of rain tomorrow in a given region over a given time period, or the maximum temperature will be 14ºC. Because these types of numbers are frequently a part of the final product for the forecaster (after a great deal of work), it seems necessary that the tools used to make forecasts would show or contain a great deal of quantitative information. Indeed, some forecasting tools are built to help the forecaster find the "best" computer model and then extract specific numeric information from that model.

But forecasters do not reason solely, or even primarily, with numbers as they try to understand and predict the weather. Forecasters glean large amounts of information from many weather visualizations and then combine that information inside their heads. An experienced weather forecaster (the type we are concerned with in this chapter) is able to create a mental model of the current atmospheric dynamics and project the likely future weather (Lowe, 1994; Perby, 1989; Trafton et al., 2000). The mental model has a significant qualitative aspect manifested as imagery, but is also "driven" by an understanding of the principles of atmospheric dynamics. Some forecasters, those who grew up on the traditional technology of hand chart work, report that their mental images are like animated charts populated with such graphic elements as isolines and wind barbs. Others report visualization of air masses and air mass interactions. Indeed, the notion of a mental model has for some time been quite familiar to the meteorology community (see Chisholm, Jackson, Niedzielski, Schechter, & Ivaldi, 1983; Doswell & Maddox, 1986) because they have to distinguish forecaster understanding ("conceptual models") from the outputs of the computer models.

The role of mental models has been demonstrated in a series of innovative experiments by Ric Lowe of Curtin University, investigating how novices and experts perceive and conceptualize the sets of meteorological markings that comprise weather charts. In Lowe's research, college student participants and weather experts carried out various tasks that required them to physically manipulate or generate markings from a given weather map. In a task in which people had to group map elements and explain the groupings, meteorologists' groupings involved the division of the map into a northern chunk and a southern chunk, which corresponds with the quite different meteorological influences that operate for these two halves of the Australian continent. Next, weather map markings were organized according to large-scale patterns that corresponded to the location of zones of regional meteorological significance. In contrast, the novices' groupings divided the map into eastern and western chunks on the basis of groups of figurally similar elements that happened to be in close proximity. Such subdivision has no real meteorological foundation.

In another task, participants were shown a map with an extended and unfilled perimeter, and had to attempt to extend the markings in the map. As well as producing significantly fewer markings in the extended region, the novices' markings appeared to have been derived quite directly from the graphic characteristics of the existing original markings by extrapolation or interpolation (e.g., turning simple curves into closed figures, continuation of existing patterns). In contrast, the meteorologists were operating in accordance with superordinate constraints involving a variety of external relations that integrated the original map area with the wider meteorological context. The resulting patterns in markings suggested the progressive clustering of lower level weather map elements into high-level composite structures that correspond to meteorologically significant features and systems of wider spatio-temporal significance.

In the task involving copying weather maps, the meteorologists began by drawing the major meteorological features. The second stage was then to pass through the map again in order to fill in subsidiary elements around this framework. In contrast, the novices tended to make a continuous pass around the map, filling in all elements they could remember in each region as they progressed, influenced primarily by the figural similarity of elements and their spatial proximity.

An especially interesting finding from a task involving the recall of maps was that the meteorologists' recall of the number of barbs on a frontal line was actually worse than that of the novices. This was because the meteorologists were concerned with the meteorologically important aspect of the cold-front symbol (the cold-front line itself) while glossing over the more "optional" aspect of the symbol (the particular number of barbs on the line).

In another task, participants attempted to predict future weather on the basis of what was shown in a map. For the nonmeteorologists, markings on the forecast maps could be largely accounted for as the results of simple graphic manipulations of the original markings; that is, they tended to move markings *en masse* from west to east without regard to meteorological dynamics. In contrast, the meteorologists' predictions showed a much greater differentiation in the way the various markings on the map were treated. Rather than moving markings *en masse,* new markings were added. This shows that meteorologists' mental representation of a weather map extends into the surrounding meteorological context.

In general, novices construct limited mental models that are insufficiently constrained, lack a principled hierarchical structure, and provide an ineffective basis for interpretation or memory. A major weakness of their mental models was the apparent lack of information available regarding the dynamics of weather systems. To quote Lowe (2000), "The expert's mental model would be of a particular meteorological situation in the real world, not merely a snapshot or image of a set of graphic elements arranged on a page" (pp. 187–188).

The consistent pattern of findings suggested a training intervention based on a new display of weather chart data. Animations were developed that portrayed temporal changes that occur across a sequence of weather maps, the idea being that animations would empower novices to develop richer mental models that would include or provide necessary dynamic information. But when novices worked with the animations, Lowe (2000) got a surprise:

> Animated material itself introduces perceptual and cognitive processing factors that may actually work against the development of a high quality mental model. . . . When the information extracted by novices was examined, it was found that they were highly selective in their approach, tending to extract material that was perceptually conspicuous, rather than thematically relevant to the domain of meteorology. . . . [For example] for highly mobile features such as high pressure cells, trajectory information was extracted while information about internal changes to the form of the feature tended to be lacking. There is clearly more research required to tease out the complexities involved in addressing ways to help meteorological novices become more adept at weather map interpretation. In particular, we need to know more about the ways in which they interact with both static and dynamic displays. (p. 205)

This finding captures the motivation for the research that we report in this chapter.

A sample display that forecasters use is shown in Figure 15.2. This shows 700-millibar heights, winds, and temperatures, and involves both a graphical (i.e., map) format and individual data points (i.e., color-coded wind

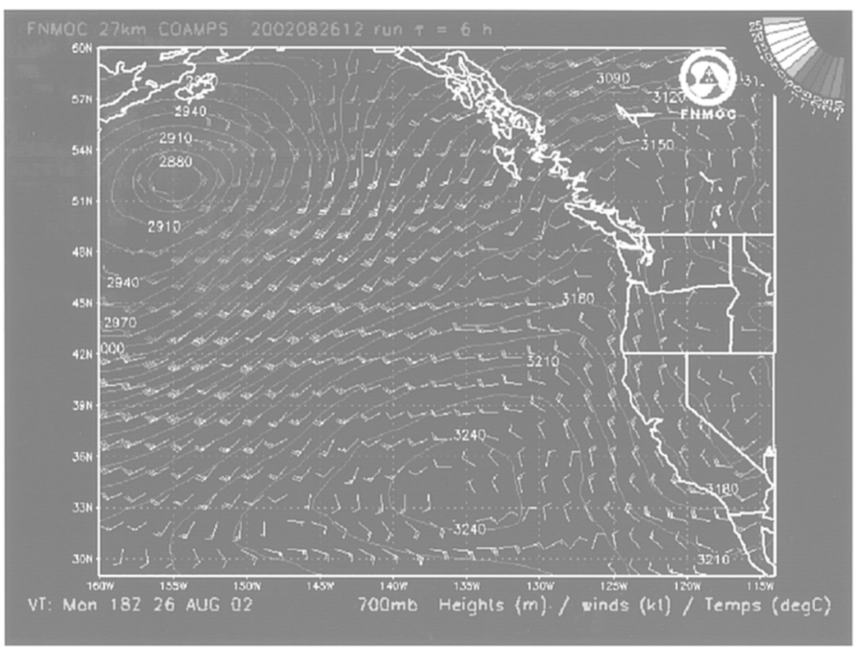

FIG. 15.2. A typical visualization that meteorologists use. Wind speed and wind direction are shown by the wind barbs; temperature is color coded according to the legend on the upper right, and equal pressure is connected by the lines (isobars) connecting the same values. The actual displays rely heavily on color, which we could not reproduce here.

barbs). Meteorological charts of this kind present a selective and somewhat decontextualized view of the atmosphere, and depict some information that is beyond direct or everyday experience. For instance, many weather charts show fronts, which are boundaries between air masses projected onto the Earth's surface. In actuality, air mass boundaries have complex shapes as a function of geopotential height (e.g., they are intersecting and interacting "blobs" of air). "Fronts" are hypotheticals. At higher levels in the atmosphere, meteorologists refer to "troughs," "ridges," and "domes":

> In weather maps these are indicated not by isolated graphical features but rather by patterning. . . . Minor local convolutions that are echoed across a series of adjacent isobars indicate the presence of a meteorologically significant feature. However, this subtle patterning of isobars can be obscured to a large extent by their visually distracting context, and so these features are likely to be overlooked unless given special attention. (Lowe, 2000, p. 189)

Hoffman (1991) and Trafton et al. (2000) have affirmed Lowe's (2000) findings: Expert forecasters do more than simply read off information from the charts or computer model outputs; they go through a process. They begin by forming an idea of the "big picture" of atmospheric dynamics, often by perceiving primarily qualitative information (e.g., "The wind is fast over San Diego" or "This low seems smaller than I thought it would be"); they continuously refine their mental model; and they rely heavily on their mental model to generate a forecast including numeric values (e.g., "The wind speed over San Diego at 500 mb will be 42 knots"). Experienced meteorologists use their mental model as their primary source of *information* (rather than copying *data* directly from the best or a favorite visualization or computer model output). Thus, simply showing a complex visualization, expecting a user to extract the necessary information, and to be finished is an oversimplification of how complex visualizations are used. Figure 15.3 shows a macrocognitive model of the forecaster reasoning process.

VISUALIZATION SUGGESTIONS

In the case of weather forecasting, the application of the Lewis and Clark Challenge and the Sacagawea Challenge is a bit tricky because forecasters must see quantitative information, but they also reason on the basis of qualitative information and their mental models. Thus, the best kind of external representation might be one that emphasizes the qualitative aspects of the data, but where quantitative information can also be immediately perceived. One good example of this is the wind-barb glyph, shown in Figure 15.2. A wind barb shows both wind speed (by the number and length of the

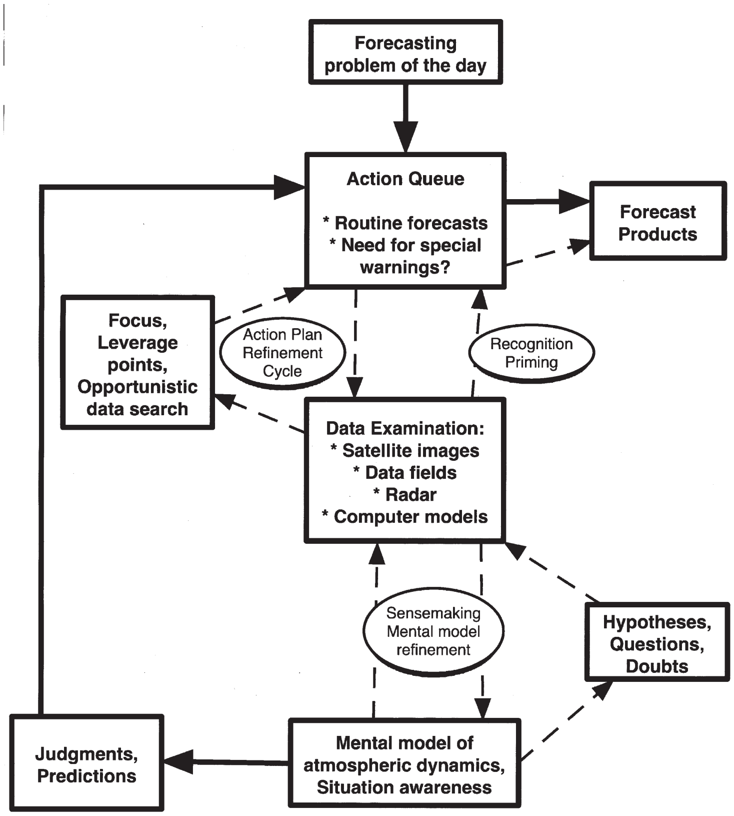

FIG. 15.3. A macrocognitive model of how expert weather forecasters predict the weather (based on Trafton et al., 2000). This diagram includes all of the fundamental reasoning processes that experts engage in: recognition-primed decision making, the action plan refinement cycle, the situation awareness cycle, and mental-model formation and refinement. In the course of cycling through multiple cycles of situation awareness and mental-model refinement, the forecaster will iterate between making products and gathering data.

barbs) and direction (by the direction of the major line). Though collapsing information in a way that might be regarded as efficient, this comes at a cost of legibility and display clutter. On the other hand, because individual wind barbs tend to cluster in ways suggestive of atmospheric dynamics (see the swirl patterns in the upper left and lower middle of Figure 15.2), the display allows the forecaster to see the qualitative aspects of the wind (areas of increasing wind speed, the formation of lows, cyclonic winds, etc.). A wind barb has the added benefit of allowing a forecaster to extract quantitative information from the glyph (e.g., each long barb is 10 knots; each short barb is 5 knots). Creating additional iconic representations that are primarily qualitative but allow quantitative information to be easily extracted from them should enhance visualization comprehension and usage.

What kinds of mental operations do weather forecasters use when predicting the weather and when working with these complex visualizations? Referring to Figure 15.3, we already know that they can make decisions by recognition priming. That is, weather data are perceived, and on the basis of experience, the forecaster knows immediately what the weather situation is and what actions need to be taken. We also know that forecasters rely on their mental models to formulate and test hypotheses and maintain a state of situational awareness. But what other kinds of mental operations are performed? To examine this issue, we developed a framework called *spatial transformations*.

A spatial transformation occurs when a spatial object is transformed from one mental state or location into another mental state or location. Spatial transformations occur in a mental representation that is an analogue of physical space and are frequently part of a problem-solving process. Furthermore, they can be performed purely mentally (e.g., purely within spatial working memory or a mental image) or "on top of" an existing visualization (e.g., a computer-generated image). Spatial transformations may be used in all types of visual-spatial tasks, and thus represent a general problem-solving strategy in this area.

There are many types of spatial transformations: creating a mental image, modifying that mental image by adding or deleting features, mental rotation (Shepard & Metzler, 1971), mentally moving an object, animating a static image (Bogacz & Trafton, 2005; Hegarty, 1992), making comparisons between different views (Kosslyn, Sukel, & Bly, 1999; Trafton, Trickett, & Mintz, 2005), and any other mental operation that transforms a spatial object from one state or location into another.

Because the mental models that meteorologists use to reason about the weather are dynamic and have a strong spatial component, it is not surprising that many spatial transformations are applied while making a forecast. We examined the type and amount of spatial transformations as expert scientists analyzed their data. We found that by far the most common spatial transformation is a comparison: Experts frequently compared and contrasted two different visualizations or compared their qualitative mental model to a visualization (Trafton et al., 2005; see also chap. 14, this volume). In fact, when the scientists used complex scientific visualizations, they performed almost twice as many comparisons as all the other spatial transformations combined. Similarly, each data field or visualization they examined was compared to others, up to four comparisons per visualization. This shows that comparisons are frequent, suggesting that they are extremely important to the forecasting process itself. Table 15.1 shows sam-

TABLE 15.1
Sample Comparisons That Forecasters Made
While Examining Meteorological Displays

Utterance by Forecaster
Location of the low, pretty similar to what, uh, NOGAPS has.
AVN at first guess looks like it has a better handle on the weather than NOGAPS.
I can't believe that radar shows precipitation, but I can't really see anything on the satellite picture.
Fleet numeric COAMPS seem to have a better handle on it than what I've seen based on the 27 kilometer.
18Z (18 Zulu or GMT time) still has a lot of moisture over the area. [At 21Z] there is not a whole lot of precipitation, and then, 00Z has a lot of moisture over the area just off the coast.
And uh, Doppler radar currently showing the precipitation a little bit closer [than I expected].
[This visualization] has a little bit more precipitation than No Gaps [or] AVN.

Note. NOGAPS, COAMPS, and AVN are computer-forecasting models.

ples of what different forecasters said as they were viewing different meteorological visualizations and making comparisons. Notice the very frequent comparisons both between and within displays that are made by the forecaster.

SUPPORT FOR MENTAL-MODEL FORMATION AND REFINEMENT

Having identified the mental model as a linchpin phenomenon in forecaster reasoning, it seems clear that what might be of use to forecasters would be a graphical tool that supports them in constructing a depiction of their four-dimensional mental model (see Hoffman, Detweiler, Lipton, & Conway, 1993). It is certainly within the reach of computer science and artificial intelligence to build a tool that might support the forecaster, for instance, in defining objects or regions within satellite images, radar, or other data types, grabbing them as graphic objects, dragging them onto a window, progressively building up a dynamic or runnable simulacrum of their understanding of atmospheric dynamics—a sketchpad to represent their mental model and spatio-temporal projections from it. Current weather information-processing workstations do not support such an activity.

Having identified the most common spatial transformation, we could also build systems that support the comparison process. Unfortunately, other than an overlay capability there is very little support for forecasters making comparisons between different weather models or to determine how well a weather model matches to a satellite image. Because comparisons are such a vital part of the forecasting process, it makes sense to support them. Specifically, there should be ways of providing forecasters with simpler ways of comparing weather models, and comparing weather models to satellite images, radar, and so forth.

With regard to both of these suggestions, there is a host of theoretical and practical research issues in the cognitive science of weather forecasting that have not yet been addressed. For example, when is it best to provide side-by-side visualizations and when is it best to provide intelligent overlays? If overlaid, how should transparency be handled? One of the difficult technical issues includes time syncing and geo-referencing the visualizations with each other so that comparisons are facilitated; it won't help the forecaster if she or he has to mentally animate one visualization to get it in sync with the other.

SUPPORT FOR THE INTERPRETATION OF QUANTITATIVE DATA

The previous section examined the spatial and reasoning processes in which forecasters engage. This section examines the perceptual processes that forecasters use to comprehend a weather visualization and extract information from it.

Most of the visualizations that forecasters examine have many variables that are used to represent upwards of tens of thousands of data points (see Hoffman, 1991, 1997). One obvious question for the designer is how to represent such an amount of data in a single visualization. There are several standard ways of accomplishing this task: color-coding variables (such as temperature), using various forms of isolines (isotachs, isobars, isotherms, etc.), and using glyphs that can combine information (as do wind barbs). All these can be used to "compress" data. However, some of these graphical tricks force the forecaster to ferret out the needed information from the mass of data. For example, one might have to interpolate between isobars or compare colors at different locations to determine which area is warmer. Such operations have to be deliberative, and can be effortful even for experienced forecasters. If we understood how forecasters perceive meaning from these types of displays, we might be able to build better ones.

Trafton et al. (2002) examined experienced forecasters' eye movements as they inspected meteorological visualizations. They found that interpolating between isobars was about twice as difficult as reading information directly off of a data chart. In the same study, they examined how forecasters extracted information by relying on legends. Extracting information from a legend can be difficult for several reasons, some cognitive, some perceptual, and some dealing primarily with the interface itself. First, the legend itself may be small or otherwise hard to find (see the temperature legend in the upper right of Figure 15.2). Second, the colors can be difficult to differentiate. Third, meteorological visualizations still tend to rely on color palettes consisting solely of highly saturated primary hues. (There are a number of lingering issues of the use of color in meteorological displays. See Hoffman et al., 1993.) Fourth, sometimes the legend is not labeled, making it unclear which of several possible variables it refers to. These factors can make finding the right match of color-to data value quite difficult.

These difficulties can manifest themselves in a variety of ways. For example, forecasters may examine different legends (if there is more than one on a visualization) or miss-guess what a legend refers to if a legend is not labeled. A forecaster may also spend an inordinate amount of time searching for the exact color if the colors are not differentiated well. Using the method of eye tracking, Trafton et al. (2002) found evidence that forecasters do all of these things when a legend is not well defined, labeled, or colored.

Trafton et al. (2002) presented experienced forecasters visualizations such as that in Figure 15.4. These visualizations all came from a forecasting Web site that was familiar to the forecasters. The researchers asked forecasters to extract specific information from these graphs while their eye movements were recorded with an eye tracker. The eye tracker allowed us to record exactly what they were looking at with a temporal resolution of 4 milliseconds. Figure 15.4 shows the relatively large amount of time and number of eye movements (saccades) one forecaster used when trying to determine what the temperature was in Pittsburgh. This back-and-forth eye movement pattern is representative of how forecasters read the legend. It shows that it took a while for the forecaster to find the right color that matched the color at Pittsburgh. Of course, these eye movements are quite fast, but over many visualizations, the time and effort invested in performing color interpretation can become substantial.

Similarly, Figure 15.5 shows a less skilled forecaster trying to figure out what the legend is showing: relative humidity or geopotential height. Notice that because the legend is unlabeled, the forecaster searches all the text in the display in the hope of determining what the legend represents. This forecaster may not have had a great deal of familiarity with this visualization, but even if she or he had, a simple label would have simplified the forecaster's search.

If we apply the Lewis and Clark Challenge and the Sacagawea Challenge to legend design, we can make some suggestions that should improve the readability of these visualizations. First, even experienced graph readers have problems color matching on legends. Most weather visualizations do not show an extremely wide range in temperature, so colors could be sepa-

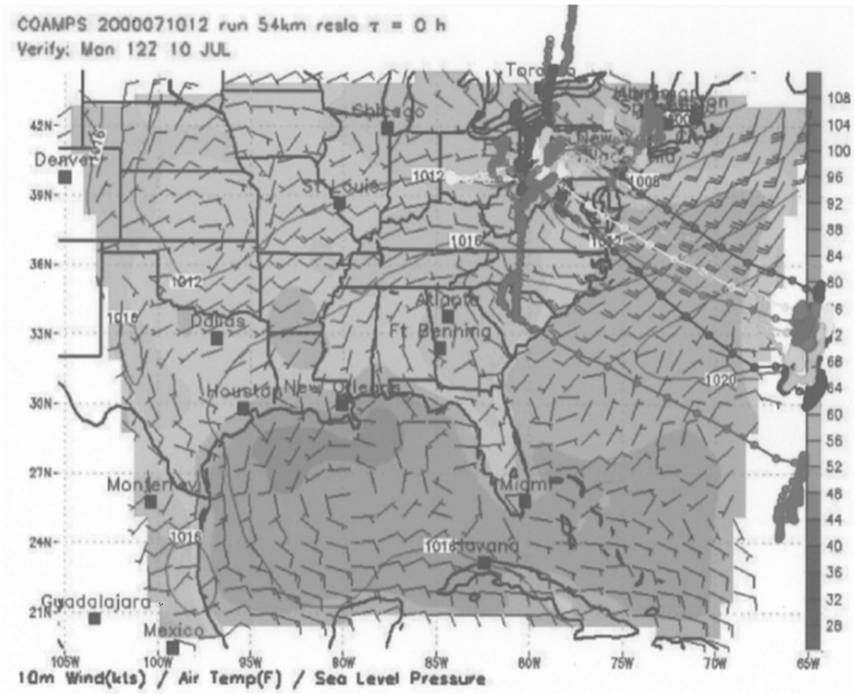

FIG. 15.4. An example visualization that was used in the study of the eye movements of weather forecasters. The forecaster was asked, "What is the temperature at Pittsburgh, Pennsylvania?" The colored lines that one can see going between the legend and the map show the eye movement tracks. The dots are 4 milliseconds apart and the changes in eye track color represent every 2 seconds, starting at Pittsburgh (obscured by the large eye-tracking trace blob). We show only some of this forecaster's eye movements for clarity.

rated more within the legend. There may still be color match problems, but the color match should be between red and blue rather than between orange-red and red-orange (see Hoffman et al., 1993). Second, we can label each legend. This is perhaps an obvious point, but it is one that even high-traffic Web sites sometimes ignore. Many newspapers, magazines, and even scientific journals do not enforce this rule. By understanding the kinds of processes that people engage in to extract information from a visualization (color matching, search of unknown information), we can attempt to prevent or ameliorate many of these problems and improve the visualizations.

An excellent example of the role that human-centering considerations can play in display design is the recent work of Lloyd Treinish and Bernice Rogowitz at IBM (e.g., Rogowitz & Treinish, 1996; Treinish, 1994; Treinish & Rothfusz, 1997). They have created a rule-based advisory tool for the

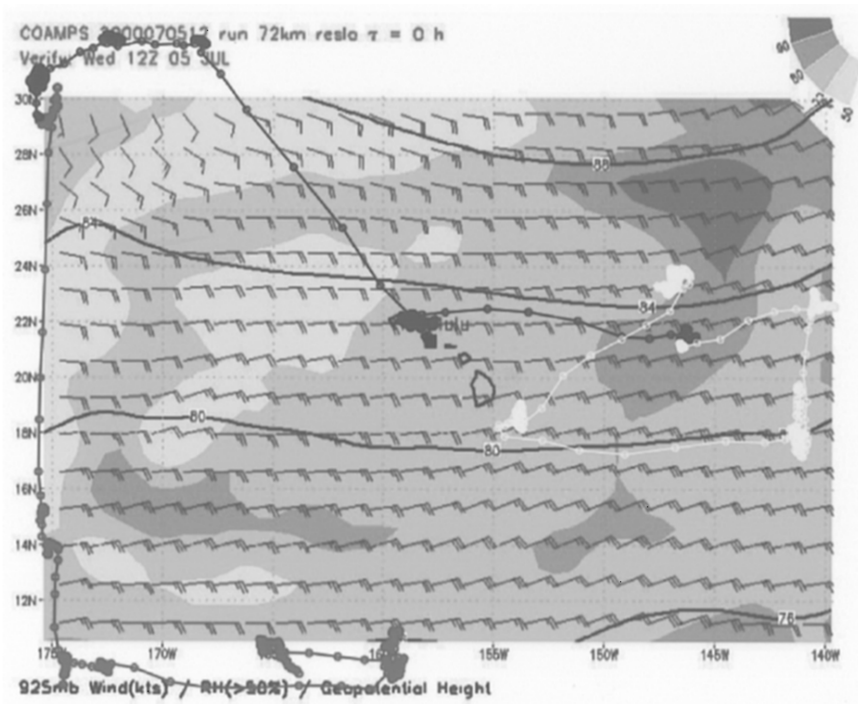

FIG. 15.5. Eye movement tracks for a forecaster attempting to answer the question, "What is the relative humidity at Honolulu?" The legend in the upper right-hand corner shows humidity, but because it is unlabeled, the forecaster searches the text and the axes to attempt to determine whether it is the isolines (solid black lines) or the legend that refers to the relative humidity.

specification of appropriate color-to-data mappings depending on whether the goal of visualization is exploration or presentation. In addition, they have relied on a human-centered strategy for visualization based on the need to preserve the fidelity of the original data, and the need to take into account known facts about human perception and cognition. Specifically, they have developed guidelines for how to collapse multiple variables and data types into individual displays and guidelines to support the user in defining coordinate systems onto which data may registered in space and time. One of their perspectival displays portrays horizontal winds (using a color palette of saturation shades of violet), relative humidity (using saturation shades of brown), surface temperature overlaid on the base map (using a two-tone palette of saturation shades of blue and green-blue), and air pressure (indicated in a semi-transparent vertical plane using saturation shades of blue-violet and green). Also depicted are three-dimensional cloud structures. For all of their graphic products, the use of perspective, depth pseudoplanes, and animation permits the perceptual discrimination of the multiple variables. (Images can be viewed at http://www.research.ibm.com/people/l/lloydt/).

CONCLUSION

In one project in which one of us was involved, the sponsor wanted to create displays to enable nonforecasters to understand the uncertainty of various types of weather data. The researchers who worked at the sponsor's laboratory assumed that these nonforecaster users would benefit by seeing the same sorts of displays that forecasters use to display data uncertainty. We knew intuitively that this might not be the best way to approach the issue. "Meaning" is always relative to the person who is doing the comprehending. Yet, it seems too easy for people, including smart, well-intentioned people, to create systems and displays that force people to work out of context. *The road to user-hostile systems is paved with user-centered intentions* (Woods & Dekker, 2001).

How might we go about improving the forecasting process, and in particular, helping forecasters stay "in context" by creating better displays? Research on the cognition of expert forecasters, compared with novices, has revealed a great deal of useful knowledge concerning the forecasting process and what is required for forecasting expertise. This knowledge can be leveraged in the application human-centering principles. We have discussed two principles of HCC, and have suggested how they might be used to improve existing meteorological products and systems. There is also a need for innovation and revolutionary redesign, especially in the creation of systems that support the forecaster in creating a graphical depiction of their own mental models of atmospheric dynamics.

ACKNOWLEDGMENTS

This research was supported in part by grants N0001400WX20844 from the Office of Naval Research and N0001400WX40002 to Greg Trafton from the Office of Naval Research. The views and conclusions contained in this document are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Navy. Thanks to Jim Ballas for comments on an earlier draft.

REFERENCES

Bogacz, S., & Trafton, J. G. (2005). Understanding dynamic and static displays: using images to reason dynamically. *Cognitive Systems Research, 6*(4), 312–319.

Chisholm, D. A., Jackson, A. J., Niedzielski, M. E., Schechter, R., & Ivaldi, C. F. (1983). *The use of interactive graphics processing in short-range terminal weather forecasting: An initial assessment* (Report No. AFGL-83-0093). Hanscom AFB, MA: U.S. Air Force Geophysics Laboratory.

Crandall, B., Klein, G., & Hoffman, R. (2006). *Minds at work: A practitioner's guide to cognitive task analysis*. Cambridge, MA: MIT Press.

Doswell, C. A., & Maddox, R. A. (1986). The role of diagnosis in weather forecasting. In *Proceedings of the 11th Conference on Weather Forecasting and Analysis* (pp. 177–182). Boston: American Meteorological Society.

Endsley, M., & Hoffman, R. R. (2002, November/December). The Sacagawea principle. *IEEE Intelligent Systems, 17*(6), 80–85.

Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*(5), 1084–1102.

Hoffman, R. R. (1990). Remote perceiving: A step toward a unified science of remote sensing. *Geocarto International, 5,* 3–13.

Hoffman, R. R. (1991). Human factors psychology in the support of forecasting: The design of advanced meteorological workstations. *Weather and Forecasting, 6,* 98–110.

Hoffman, R. H. (1997, February). *Human factors in meteorology*. Paper presented at the Short Course on Human Factors Applied to Graphical User Interface Design, held at the 77th annual meeting of the American Meteorological Society, Los Angeles.

Hoffman, R. R., & Conway, J. (1990). Psychological factors in remote sensing: A review of recent research. *Geocarto International, 4,* 3–22.

Hoffman, R. R., Detweiler, M. A., Lipton, K., & Conway, J. A. (1993). Considerations in the use of color in meteorological displays. *Weather and Forecasting, 8,* 505–518.

Hoffman, R. R., & Lintern, G. (2006). Eliciting and representing the knowledge of experts. In K. A. Ericsson, N. Charness, P. Feltovich, & R. Hoffman (Eds.), *Cambridge handbook on expertise and expert performance* (pp. 203–222). New York: Cambridge University Press.

Hoffman, R. R. & Markman, A. B. (Eds.). (2001). *Human factors in the interpretation of remote sensing imagery*. Boca Raton, FL: Lewis.

Hoffman, R., Trafton, G., & Roebber, P. (2006). *Minding the weather: How expert forecasters think*. Cambridge, MA: MIT Press.

Klein, G. A., & Hoffman, R. R. (1992). Seeing the invisible: Perceptual-cognitive aspects of expertise. In M. Rabinowitz (Ed.), *Cognitive science foundations of instruction* (pp. 203–226). Mahwah, NJ: Lawrence Erlbaum Associates.

Kosslyn, S. M., Sukel, K. E., & Bly, B. M. (1999). Squinting with the mind's eye: Effects of stimulus resolution on imaginal and perceptual comparisons. *Memory and Cognition, 27*(2), 276–287.

Lowe, R. K. (1994). Selectivity in diagrams: Reading beyond the lines. *Educational Psychology, 14,* 467–491.

Lowe, R. (2000). Components of expertise in the perception and interpretation of meteorological charts. In R. R. Hoffman & A. B. Markman (Eds.), *Interpreting remote sensing imagery: Human factors* (pp. 185–206). Boca Raton, FL: Lewis.

Monmonier, M. (1999). *Air apparent: How meteorologists learned to map, predict, and dramatize weather*. Chicago: University of Chicago Press.

Norman, D. (1993). *Things that make us smart: Defending human attributes in the age of the machine*. Cambridge, MA: Perseus.

Perby, M. L. (1989). Computerization and skill in local weather forecasting. In I. Josefson (Ed.), *Knowledge, skill and artificial intelligence* (pp. 39–52). Berlin: Springer-Verlag.

Rogowitz, B. E., & Treinish, L. A. (1996). How not to lie with visualization. *Computers in Physics, 10,* 268–274.

Schraagen, J.-M., Chipman, S. F., & Shalin, V. L. (Eds.). (2000). *Cognitive task analysis.* Mahwah, NJ: Lawrence Erlbaum Associates.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701–703.

Trafton, J. G., Kirschenbaum, S. S., Tsui, T. L., Miyamoto, R. T., Ballas, J. A., & Raymond, P. D. (2000). Turning pictures into numbers: Extracting and generating information from complex visualizations. *International Journal of Human–Computer Studies, 53*(5), 827–850.

Trafton, J. G., Marshall, S., Mintz, F., & Trickett, S. B. (2002). Extracting explicit and implicit information from complex visualizations. In M. Hegarty, B. Meyer & H. Narayanan (Eds.), *Diagramatic representation and inference* (pp. 206–220). Berlin: Springer-Verlag.

Trafton, J. G., Trickett, S. B., & Mintz, F. E. (2005). Connecting internal and external representations: Spatial transformations of scientific visualizations. *Foundations of Science, 10*(1), 89–106.

Treinish, L. A. (1994). Visualization of disparate data in the earth sciences. *Computers in Physics, 8,* 664–671.

Treinish, L. A., & Rothfusz, L. (1997). Three-dimensional visualization for support of operational forecasting at the 1996 Centennial Olympic Games. In *Proceedings of the Thirteenth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology* (pp. 31–34). Boston: American Meteorological Society.

Vicente, K. (2000). *Cognitive work analysis.* Mahwah, NJ: Lawrence Erlbaum Associates.

Woods, D. D., & Dekker, S. W. A. (2001). Anticipating the effects of technology change: A new era of dynamics for human factors. *Theoretical Issues in Ergonomics Science, 1,* 272–282.