



Understanding dynamic and static displays: using images to reason dynamically

Action editor: Wayne Gray

Sally Bogacz^{a,*}, J. Gregory Trafton^{b,1}

^a Clarify Concepts, 1330 New Hampshire Avenue, NW, #915, Washington DC 20036, USA

^b Naval Research Laboratory, Washington DC, USA

Received 5 September 2004; accepted 30 November 2004

Available online 8 January 2005

Abstract

We examined expert meteorologists as they created a weather forecast while working in a naturalistic environment. We examined the type of external representation they chose to examine (a static image, a sequence of static images, or a dynamic display) and the kind of information they extracted from those representations (static or dynamic). We found that even though weather is an extremely dynamic domain, expert meteorologists examined very few animations, examining primarily static images. However, meteorologists did extract large amounts of dynamic information from these static images, suggesting that they reasoned about the weather by mentally animating the static images rather than letting the software do it for them.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Representation; Images; Static; Dynamic; Animations; Meteorology; Weather forecasting; Reasoning

1. Introduction

What kind of external displays do experts use to extract dynamic information? External displays are usually either static (a diagram), or dynamic

(an animation). This question is particularly relevant for domains that have a strong spatial component to them like scientific reasoning (Schunn & Anderson, 1999; Trafton, Trickett, & Mintz, in press; Trickett & Trafton, under review), meteorology (Lowe, 1994, 1999; Trafton et al., 2000), and medical diagnosis (Lesgold et al., 1988).

Research examining the use of static and dynamic displays shows mixed results across many different domains. The problem with *static* images

* Corresponding author. Tel.: +1 202 331 9267.

E-mail addresses: sally@clarifyconcepts.com (S. Bogacz), trafton@itd.nrl.navy.mil (J.G. Trafton).

¹ Tel.: +1 202 767 3479.

is that they can impose a high working memory load when the task is to reason about a machine in motion (Hegarty & Sims, 1994; Narayanan, Suwa, & Motoda, 1994; Sims & Hegarty, 1997). However, many studies have found that *animations* by themselves do not improve performance (Byrne, Catrambone, & Stasko, 1999; Mayer & Anderson, 1991; Palmiter & Elkerton, 1993; Palmiter, Elkerton, & Baggett, 1991; Rieber, Boyce, & Assad, 1990) unless they provide more information than static images (Pane, Corbett, & John, 1996; Tversky, Morrison, & Betrancourt, 2002).

The finding that animations by themselves do not improve performance has led many researchers to question their usefulness (Palmiter & Elkerton, 1993; Pane et al., 1996), suggesting that animations should be used only in very limited situations, i.e. only when necessary and when the animation is not too difficult to use (Betrancourt & Tversky, 2000). But most of these studies have been performed in laboratory settings (e.g., Kaiser, Proffitt, Whelan, & Hecht, 1992; Palmiter et al., 1991) or use tasks specially crafted to show animated or static pictures (e.g., Pane et al., 1996; Rieber, 1991). There have been very few studies that have looked at how (or why) domain experts use animations in complex, dynamic domains. What happens, for example, when experts in a dynamic domain have a choice of whether to use static or animated images? What type of images do they look at, and what type of information do they extract from these images? How do they use this information to help them solve problems?

Most researchers in the technology field believe that animations are an important tool that can help us to understand complex domains. For this reason, animations have been used in recent years to teach procedures in HCI (Palmiter & Elkerton, 1993), to teach computer science algorithms (Byrne et al., 1999) to teach how something works (Mayer & Anderson, 1991; Pane et al., 1996), and to understand other complex dynamic systems, like the weather (Lowe, 1999). Indeed, the prevalent feeling in this body of literature is that animations *should* be better than static images because, by a principle of congruence, animations should be a natural medium for conveying infor-

mation about change, just as graphics are a natural way for conveying information about space (Tversky et al., 2002).

One domain that seems tailor made for the use of animations is meteorological forecasting. The forecaster has to determine the dynamics of past and current weather and predict what (if anything) will change in the future. So animations should be useful to forecasters because the domain they work in forces them to explicitly consider the relation between directional movements over time and space (Rieber, 1990; Rieber & Kini, 1991), as well as real-time changes (Tversky et al., 2002).

Previous studies that have examined the forecasting process have shown that meteorologists use a wide range of information to do their job: static images, observations of the data, satellite pictures, computational weather models, display loops, textual information and dopplar radar (Hoffman, 1991; Trafton et al., 2000). Most weather web sites present information in the form of both static and animated displays.

In the study discussed below, participants were skilled Naval meteorologists. One of their primary sources for weather displays and other information was the Fleet Numerical Meteorology and Oceanography Center (FNMOC) website (www.fnmoc.navy.mil). This website had a portal with links to different displays, as shown in Fig. 1.

Using this portal, the forecaster could view a static image by clicking on the button marked “000” in the row marked “TAU”. He would see something similar to Fig. 2, which shows relative vorticity (or atmospheric swirl) for the present time.

Fig. 2 shows a NOGAPS weather model output for 4/29/2002. τ is TAU which refers to the time in the future for which this weather model will be valid. The zero in this case means that it is a “prediction” of the current weather. The display shows pressure at 500 mb and the amount of relative vorticity or wind spin. The original is in color.

Alternatively, he could view a series of images showing vorticity at different times (present time, 12 h into the future, etc.), by clicking on the button marked “all” to the left under “TAU”, and would see something similar to Fig. 3.

FNMOC WXMAP		Model: NOGAPS					Area: North America				DTG: 2002082912				
TAU	Time (now) and hours in the future	000	012	024	036	048	060	072	084	096	108	120	132	144	Loop
<u>all</u>	500 hPa Heights [m] and Rel. Vort [10^{-5} s^{-1}]	●	●	●	●	●	●	●	●	●	●	●	●	●	●

Fig. 1. An example of a FNMOC portal with choices of different types of display.

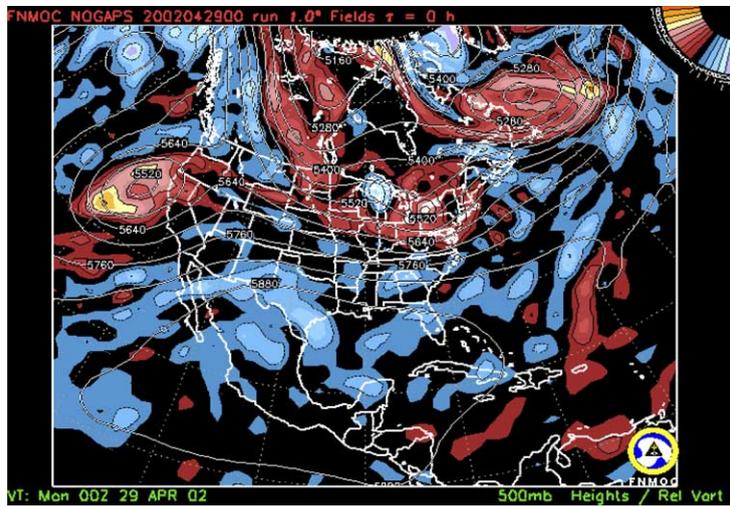


Fig. 2. An example of a picture: FNMOC NOGAPS 2002042900 $t = 0$ h forecast of 500 hPa Heights (m) and Rel. Vort (10^{-5} s^{-1}).



Fig. 3. An example of a sequence of images: FNMOC NOGAPS 2002042900 500 hPa Heights (m) and Rel. Vort (10^{-5} s^{-1}) maps for all times.

Forecasters could scroll through these figures after they appeared in a browser or other window. Sequences were always a series of static pictures of the weather that started at a specific time and projected forward in equal time increments (typically 6 or 12 h) into the future.

Finally, forecasters could view an animation of vorticity changing over time by clicking on the button marked “Loop”, where he would see something similar to Fig. 4.

So in their normal work environment, forecasters had a choice about what type of display to look at. In this study, we examine what happens when forecasters can choose what to look at, by focusing on expert meteorologists as they create a weather report in naturalistic surroundings.

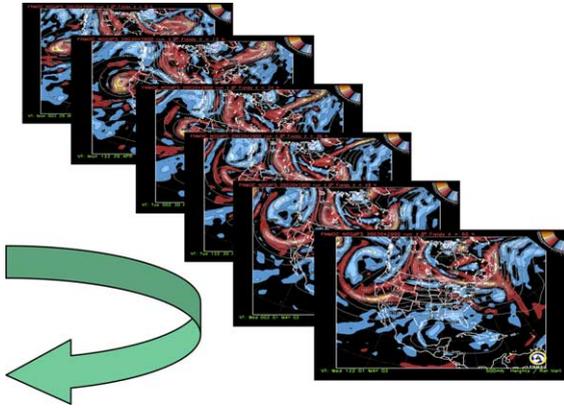


Fig. 4. An example of an animation: FNMOC NOGAPS 2002042900 500 hPa Heights (m) and Rel. Vort (10^{-5} s^{-1}) maps for Java Movie. The green (gray in print) arrow indicates that this sequence of images was put into a repeatedly cycling loop.

2. Method

The experiment examined how expert meteorologists made forecasts in a complex work environment that simulated the situation on board ship.

2.1. Participants

All forecasters were Naval or Marine Corps forecasters and forecasters-in-training. The participants were representative of the range of expertise and training typically found for Naval METOC (METeological and OCEanographic) forecasters.²

Two individuals participated in each session, playing the role of local forecasters on the ship. They worked in pairs, with the more experienced individual acting as lead forecaster and the less experienced as technician or forecaster-in-training. Each forecaster provided talk-aloud protocols (Ericsson & Simon, 1993). The data for this study consists of the utterances made by the expert forecaster of each of these pairs ($N = 7$), who had an average of 11 years of forecasting experience. Each forecaster had prepared over 100 weather forecasting briefs the previous year.

² All forecasters were men, so male pronouns will be used throughout this paper to refer to them.

2.2. Apparatus and setting

The experimental sessions took place in a large room. On one side of the room, two Windows-NT workstations were arranged side by side for the forecaster and technician. Video recorders were positioned to capture the forecasters' and technicians' computer screens, with a third one positioned to capture the interactions between the two forecasters.

Two subject-matter experts (SME) stood behind the forecasters and took notes about what the forecasters were doing. On the opposite side of the room were three workstations that were used as the regional center under the supervision of a senior METOC officer. All communications between the simulated shipboard and the simulated Regional Center were carried out via a chat tool called IRC Chat that all participants were familiar with.

2.3. Forecasting tools

The forecasters frequently used two programs to visualize the weather data: Joint METOC Viewer (JMV) and a World Wide Web browser. JMV allows weather model data to be visualized in various ways: as a still, single image, as a sequence of images showing the same meteorological information separated by time (e.g., 0, 6, 12, 18, 24 h into the future) and as an animation which consisted of the sequence of images put into a loop that cycled, repetitively, as shown in Figs. 2–4. The World Wide Web browser allows forecasters to view satellite images as well as METAR (aviation routine weather report) and terminal aerodrome forecast (TAF) text from various websites. Satellite images could also be static or dynamic.

It should be noted that information that is displayed by all major meteorological web sites (e.g., FNMOC, NOAA) shows identical information in different ways: static pictures, sequences, or animated sequences of pictures. So all animations that a forecaster could look at had the same information as their *static* or *sequence* counterparts. What most systems do in order to animate meteorological information is to generate the individual consecutive displays and animate them by presenting

them quickly, as though the user were flipping through a flip-chart. This method is used to display almost all time-sequence meteorological products, including global weather, local weather and satellite imagery.

2.4. Procedure

The task was to prepare a written brief for an airplane flown from an aircraft carrier to a destination 12 h in the future (the destination was Whidbey Island, Washington State). The brief was to cover the entire round trip and the forecasters were asked to provide specific weather information for departure, en-route, destination and alternate airfields. In order to do this, forecasters had to determine detailed qualitative and quantitative information about the weather conditions. This task was a very familiar one for the forecasters.

Each session began with a description of the task by an experienced Navy forecaster, giving destination, times and other information. Forecasters created briefs using PowerPoint or wrote the information down on paper. The forecaster served as leader during the session, requesting information from the technician as needed.

Forecasters were given 2 h to complete their forecast and brief. Everyone completed the task within the time allotted. If time allowed, the forecaster presented his brief at the end of the session. All sessions were concluded with a debriefing during which the experimenter had an opportunity to ask questions and the participants had an opportunity to give feedback to the experimenters.

2.5. Coding scheme

The forecasters' utterances were transcribed and coded, and the external representations they used were noted by examining the sessions on videotape with an audio soundtrack.

2.5.1. Coding of external representations

Forecasters used many data sources. Satellite images and JMV displays were very commonly used, but forecasters also examined Skew-Ts (temperature profiles at a particular location), METAR and TAF text as well as several other

Table 1
Sample coding scheme of external representations

Category	Example
Picture	COAMPS/San Diego/18Z (6 p.m. Zulu time is 6 p.m. Greenwich mean time or GMT)
Sequence	NOGAPS/Whidbey Is./0Z (midnight GMT), 6Z (6 a.m. GMT), 12Z, 18Z, 24Z
Animation	Intellicast/Satellite loop
Text	Port Angeles METAR text page

displays. Each time the forecaster examined a display, it was coded (see Table 1).

The external representations were categorized as a *picture* (a still weather map or satellite image), a *sequence* (a sequence of weather maps or satellite images separated by equal time intervals), an *animation* (the sequence of weather maps or satellite images put into a continuously cycling loop) or *text* (METAR and TAF text reports) (see Figs. 2–4).

2.5.2. Utterance coding

Utterances were categorized for weather-related extractions. Remarks made by the forecaster about goals, military matters or astronomical observations were excluded unless they contained information about the weather. In addition, utterances that were not related to information gleaned from displays shown on the forecaster's screen, such as questions that tutored the technician ("Did you interpret it correctly off the TAF?") or extraneous information (e.g., how to find a NOGAPS product on the Internet) were also excluded. Each of the remaining utterances was coded for its dynamic content.

2.5.3. Dynamic and static utterances

Utterances were coded for their dynamic or static content. Static utterances described the weather at one location at a specific time. Dynamic utterances captured change across time or space. These included utterances where change was implied rather than specifically stated: For example, "It's going to be valid after 13:45Z" or "So everything from like 12Z on put a broken on at about 500 feet". Both of these statements are describing conditions for a particular period of time, implying that outside that time-frame, conditions change. Examples can be found in Table 2.

Table 2
Sample coding scheme of static and dynamic weather utterances

Category	Example utterances
Static	Whidbey has a ceiling of 4500 broken Mostly sunny afternoon, 70, 75 Seattle
Dynamic	24 h from now, definitely a lot of precipitation moving into the area There's a frontal system approaching the area producing precipitation

To ensure reliability of coding, one author coded the entire dataset and the other author coded 25% of it. The two coders agreed 87% of the time: $\kappa = .76$, $p < .001$. Disagreements were resolved by discussion.

3. Results and discussion

This section focuses on forecasters' use of animations and the implications for reasoning in a complex dynamic domain. Given a choice, do forecasters use animations? Do dynamic images help forecasters reason about the weather?

3.1. Overview

The forecasters looked at an average of 21.4 external representations each. They extracted information from these external representations an average of 5.3 times per representation. These numbers are probably an underestimate of the amount of information extracted because utterances were not coded when the forecaster was tutoring the technician, which happened fairly frequently.

3.2. External representations

Forecasters had a choice of what image to look at. There were four categories: *picture*, *sequence*, *animation* and *text*. Fig. 5 shows how often (in percentages) the forecasters looked at each type of image (see light bars).

Notice that forecasters made *almost no use* of animations. Instead, forecasters chose to look at pictures (which are *static* images) most of the time, $\chi^2(3) = 70.32$, $p < .001$, Bonferroni adjusted χ^2 significant at $p < .001$.

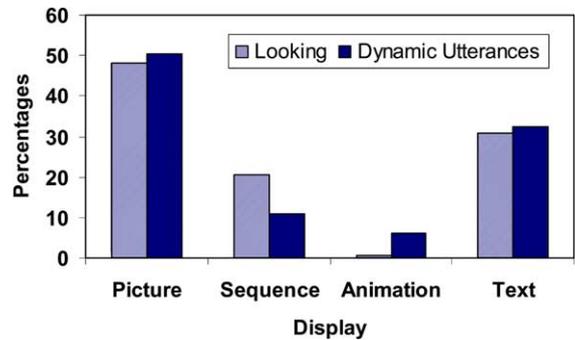


Fig. 5. The percentage of dynamic utterances and percentage of different types of display examined.

What is puzzling about this result is that animations were just as available for use as sequences, pictures or text. Does this lack of animation use mean that meteorologists were *not* thinking dynamically when forecasting? We examined this question by looking at the types of utterances (*static* or *dynamic*) that forecasters made while looking at different kinds of display.

3.3. Utterances

We found that forecasters were thinking dynamically about the weather for a substantial proportion of the time: 35% of the forecasters' utterances were dynamic. Thus forecasters made little use of animated images, but talked dynamically about the weather for over a third of the time. It could be that most of the dynamic utterances were made while examining the (relatively rare) animated images or the (slightly more common) sequences of images. However, Fig. 5 demonstrates (see dark bars) that forecasters made dynamic utterances constantly, with the rate of production of dynamic utterances being approximately proportional to display use. So because forecasters did not use animations, they also directed very few dynamic utterances towards them.

4. General discussion

When expert meteorological forecasters made weather predictions, they examined static rather than animated displays of data. Forecasters

preferred to use static displays even though there were many dynamic displays available and they knew how to use them. They used the static displays even though they were thinking dynamically about the weather. And they produced dynamic utterances at the same rate as they looked at each type of display.

These results are not surprising if you read the large body of literature showing that, in laboratory settings, there is very little advantage to using animations when they contain the same amount of information as static displays (Byrne et al., 1999; Rieber et al., 1990). But technology designers, who believe that animations are an important tool in helping us to understand complex domains, will be surprised. Animations *should* be better than static images at describing the weather, because animations show change over time (e.g., the “Principle of Congruence”; Tversky et al., 2002). Isn’t that what forecasters want to know?

Our research indicates, however, that a dynamic animation is *not* the preferred method of retrieving dynamic information about the weather. This agrees with other research showing that people do not perform better or learn more when using a dynamic display. Explanations for this lack of effect include the idea that animations impose a high workload (Lowe, 2000), a lack of knowledge about how we integrate internal and external representations (Scaife & Rogers, 1996), comprehension difficulty (Tversky et al., 2002), difficulty in perceiving (Kaiser et al., 1992) and difficulty in use (Betrancourt & Tversky, 2000).

How does an expert forecaster extract that dynamic information from static displays? Consistent with other research on static displays (e.g., Hegarty & Sims, 1994), we believe that forecasters animate static displays (or parts of the displays) mentally, and use that information to create a dynamic mental representation of atmospheric dynamics, e.g., how a front would move. Or they blend different weather models to create a dynamic mental representation of the most probable path for a weather system. Our best guess as to why our expert participants did not use animations is that these meteorologists had enough expertise to be able to memorize critical global features, but needed to see the details which are visible only in static

images. If they mentally animated the static displays, this would have given them control over the mental images they created, enabling them to “set in motion” any relevant details. It is clear, however, that not only is it possible for experts to extract dynamic information from static displays, but that it is their preferred method of doing so.

Acknowledgements

This work was supported in part by Grants N00014-00-WX-20844 and N00014-00-WX-40002 to Greg Trafton from the Office of Naval Research. We thank the forecasters who participated in this study. We also thank Susan Trickett and Nick Cassimatis for comments on an earlier draft and four anonymous reviews for comments on a later draft.

References

- Betrancourt, M., & Tversky, B. (2000). Effect of computer animation on users’ performance: a review. *Travail Humain*, 63(4), 311–329.
- Byrne, M. D., Catrambone, R., & Stasko, J. (1999). Evaluating animations as student aids in learning computer algorithms. *Computers and Education*, 33(4), 253–278.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge MA: MIT Press.
- Hegarty, M., & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory and Cognition*, 22(4), 411–430.
- Hoffman, R. R. (1991). Human factors psychology in the support of forecasting: the design of advanced meteorological workstations. *Weather and Forecasting*, 6(1), 98–110.
- Kaiser, M. K., Proffitt, D. R., Whelan, S. M., & Hecht, H. (1992). Influence of animation on dynamical judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 669–690.
- Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing X-ray pictures. In M. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise* (pp. 311–342). Hillsdale, NJ: Erlbaum.
- Lowe, R. K. (1994). Selectivity in diagrams: reading beyond the lines. *Educational Psychology*, 14, 467–491.
- Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology and Education*, XIV(2), 225–244.

- Lowe, R. K. (2000). Animation of diagrams: An aid to learning? In M. Anderson, P. Cheng, & V. Haarslev. Theory and application of diagrams. Heidelberg: Springer.
- Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual coding hypothesis. *Journal of Educational Psychology*, 83(4), 484–490.
- Narayanan, N.H., Suwa, M., & Motoda, H. (1994). A study of diagrammatic reasoning from verbal and gestural data. In *16th Annual Conference of the Cognitive Science Society*.
- Palmiter, S., & Elkerton, J. (1993). Animated demonstrations for learning procedural computer-based instruction. *Human-Computer Interaction*, 8, 193–216.
- Palmiter, S., Elkerton, J., & Baggett, P. (1991). Animated demonstrations versus written instructions for learning procedural tasks: a preliminary investigation. *International Journal of Man-Machine Studies*, 34, 687–701.
- Pane, J. F., Corbett, A. T., & John, B. E. (1996). Assessing dynamics in computer-based instruction. In *CHI '96, Vancouver, BC, Canada*.
- Rieber, L. P. (1990). Animation in computer-based instruction. *Educational Technology, Research and Development*, 38, 77–86.
- Rieber, L. P. (1991). Animation, incidental learning and continuing motivation. *Journal of Educational Psychology*, 83, 318–328.
- Rieber, L. P., Boyce, M. J., & Assad, C. (1990). The effects of computer animation on adult learning and retrieval tasks. *Journal of Computer-Based Instruction*, 17(2), 46–52.
- Rieber, L. P., & Kini, A. S. (1991). Theoretical foundations of instructional applications of computer-generated animated visuals. *Journal of Computer-Based Instruction*, 18, 83–88.
- Scaife, M., & Rogers, Y. (1996). External cognition: how do graphical representations work. *International Journal of Human-Computer Studies*, 45(2), 185–213.
- Schunn, C. D., & Anderson, J. R. (1999). The generality/specificity of expertise in scientific reasoning. *Cognitive Science*, 23(3), 337–370.
- Sims, V. K., & Hegarty, M. (1997). Mental animation in the visuospatial sketchpad: evidence from dual-task studies. *Memory and Cognition*, 25(3), 321–332.
- Trafton, J. G., Kirschenbaum, S. S., Tsui, T., 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, 827–850.
- Trafton, J. G., Trickett, S. B., & Mintz, F. (in press). Connecting internal and external representations: spatial transformations of scientific visualizations. *Foundations of Science*.
- Trickett, S. B., & Trafton, J. G. (under review). Movies in the mind: the instantiation and use of conceptual simulations in scientific reasoning.
- Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies*, 57(4), 247–262.