Design and assessment of a geographic visualization tool to support earth science learning

Mark Harrower

Alan MacEachren

Amy L. Griffin

GeoVISTA Center, The Pennsylvania State University, 302 Walker Building, University Park, PA 16802-5011, USA

mah282@psu.edu; alan@geog.psu.edu; alg207@psu.edu

ABSTRACT

This paper reports on the development and assessment of an animated and interactive geovisualization environment and on implications of this environment for learning about spatiotemporal processes. This environment, the EarthSystemVisualizer (ESV), is designed to facilitate learning about global weather. Our goals in designing ESV were (1) to evaluate two exploratory spatial data analysis (ESDA) techniques: *temporal brushing* and *temporal focusing* and (2) to determine whether an interactive geovisualization environment influenced problem solving strategies, approaches to learning, and students' ability to generate hypotheses about earth-science processes. Focus group sessions were conducted with both expert and novice users to assess an initial design for the ESV interface prior to conducting a task-based assessment of ESV use. Changes were implemented in response to the focus group results, including the redesign of a temporal legend and improved speed and direction controls. Our task-based assessment considered student reactions to components of the ESV, whether the ESV could be used by students to answer objective questions about global scale weather processes, and whether the system (particularly its focusing and brushing tools) had an impact on hypotheses generated about relationships among weather variables. The assessment revealed that focusing and brushing had little impact on students' ability to answer a series of objective questions. Performance suffered for students who were confused by the focusing and brushing tools. In fact, students without the tools performed better than did those who were confused by the tools, but students who understood the tools performed the best. We also concluded that the level of the visualization system must be well matched to the knowledge level of the user about the application domain: students who already possessed an advanced understanding of meteorology or climatology benefited less and were more critical of the system than students with an intermediate or novice level understanding.

KEYWORDS: map animation, geographic visualization, interactive cartography, cartographic design, legends, spatiotemporal analysis, user testing

INTRODUCTION

Understanding earth system processes demands a sophisticated comprehension of the temporal and spatial aspects of those processes; air temperature, for example, contains diurnal, weekly, seasonal, and inter-annual patterns that are played out differentially across the earth's surface. Trying to understand the relationships among multiple climate variables complicates this task. Visually representing such complex spatiotemporal relationships is a significant design challenge. Recent developments in Geographic Visualization (geovisualization) have adapted techniques from exploratory data analysis such as *sequencing* (Slocum 1988; Monmonier 1992; Peterson 1995), *brushing* (Monmonier 1989; Edsall and Peuquet, 1996), and *focusing* (Cook et al. 1996; MacEachren et al. 1997). In turn, these have led to the development of specific tools such as linked geographic displays (MacDougall 1992), coupled statistical-geographic representations (Monmonier 1989; Becker et al. 1988), and dynamic on-screen classification techniques (MacEachren et al. 1997). These tools provide designers with new techniques for representing georeferenced information and users with new ways of interacting with these representations. However, many questions remain unanswered about how these new tools and representational techniques influence problem solving and knowledge construction in the geographic domain.

We report here on a project directed toward developing and testing a set of space-time visualization tools designed to facilitate earth science learning and knowledge construction. Our research has two primary goals. The first is to integrate two exploratory data analysis methods (*brushing* and *focusing*) with map animation to produce a dynamic, interactive representation that represents time as both linear and cyclic. The second goal is to implement these tools in a geovisualization system that allows users to explore complex spatial and temporal aspects of multivariate continuously changing phenomena (specifically weather and climate).

To meet these goals, we built the EarthSystemVisualizer (**Figure 1**). ESV allows users to explore the spatiotemporal behavior of three aspects of global weather: land temperature, ocean temperature, and cloud cover. ESV is designed primarily as an introductory-level undergraduate educational tool. Much of the research effort in geovisualization has been directed toward experts who possess high-level knowledge of the subject matter (e.g. climatologists) and who work routinely with large and complex data sets (e.g. statisticians). We believe that exploratory data analysis techniques such as *brushing* and *focusing,* although developed for experts, are appropriate for all levels of expertise and can assist in problem solving and learning even at an introductory level.

Key objectives of this study were to determine whether ESV tools (1) are understood and used effectively to answer questions about spatial, temporal and attribute components of climate / weather patterns and processes, (2) prompt different knowledge schemata than those applied without these tools, (3) stimulate different approaches to problem solving, and ultimately, (4) lead to generation of different hypotheses about the relationship between climate variables over both space and time. Our intent was not simply to determine if these tools "work," but rather, to determine if they influence problem solving and learning and if so, how.

This research was conducted in three stages. First, we developed a prototype of ESV. Second, we used focus groups to assess and refine the initial system. Finally, formal user testing of the revised ESV was undertaken in a controlled environment.

This paper contains four sections. Section One reviews relevant literature needed to place our work in the broader context of recent developments in geovisualization. This review touches upon three themes: cartographic representations of change and map animation, the integration of exploratory data analysis (EDA) and cartographic methods, and learning from maps. Section Two introduces ESV and the rationale behind its development. Section Three discusses focus groups and what we learned from this qualitative approach. Section Four details the formal task-based user testing of ESV, the results of these tests, and insights into user learning and knowledge construction as they relates to interface design.

SETTING THE CONTEXT: AN OVERVIEW

Cartographic representations of change and Map Animation

The representation of change has been an increasingly important topic for cartographers in recent years. Szegö (1987), for example, devotes much of his text *Human Cartography* to this topic, integrating a perspective of time geography derived from Hägerstrand (1974) with cartographic representation methods for static display. Monmonier (1990) provides an excellent overview of the ways in which static maps can depict changes.

Representing change over time on a single static map is, however, a difficult problem with no ideal solution – just many alternatives that can highlight different kinds or aspects of change (see, for example, Szegö 1987; MacEachren 1995; Edsall et al. 1997). Cartographers have been interested in the potential of map animation, in particular, as a way to depict changing phenomena for approximately four decades (see Thrower 1961 for one of the first accounts of cartographic animation and Campbell and Egbert 1990 for a review of early map animation research). Much of the early work in map animation conceptualized time as linear and focused on techniques for generating map movies to represent change over time (e.g. Thrower 1961; Tobler 1970; Gould et al. 1990; Dorling and Openshaw 1992). Research in this area includes appropriate metaphors for representing change (Gersmehl 1990); choice of symbolization methods for depicting changing phenomena (MacEachren and DiBiase 1991); design of graphic scripts that organize blocks of time into a coherent story (Monmonier 1990); uses of animation in television, such as weather broadcasts (Carter 1996); and issues of interpolation needed to generate smooth animations from temporally sparse data (Acevedo and Masuoka 1997).

Although map animation research has generally emphasized a linear perspective on time, Moellering (1976) demonstrated more than two decades ago with an animation of traffic accidents the power of map animation to facilitate an understanding of cyclic phenomena. Moellering’s innovation was to depict the location and time of accidents for a composite week in frames representing 15-minute intervals containing all accidents during that time on that day of the week for the full three-year period. The result was a clear represtation of the daily cycle of accidents, with peaks during rush hour periods and troughs between those times, as well as the spatial pattern of weekday versus weekend accidents.

The potential power of animated maps lies in their ability to prompt a conceptualization of temporal continuity, thus facilitating an understanding of process rather than state. As Monmonier and Gluck (1994) found, however, viewers are often frustrated by complex changing maps that they cannot control. This was echoed by Koussoulakou and Kraak (1992) who found no difference in effectiveness between static maps and animated maps when the animated maps could not be controlled. In response to these findings, research attention has turned to manipulable animated displays. Manipulation can take many forms, from direction and pace controls (Monmonier and Gluck 1994) to an ability to zoom to a different scale of analysis (Buttenfield and Weber 1994).

Integration of EDA with cartographic representation

An important component of the overall research thrust associated with geovisualization has been the integration of EDA methods into map-based environments. Several graphical EDA methods, including *linked brushing* (Carr et al. 1987; Monmonier 1989), *focusing* (MacEachren et al. 1993), and the "grand tour" (Monmonier 1992) have been adapted and extended for application with georeferenced information. EDA methods have also been extended to deal with temporal data. Monmonier (1989) proposed the concept of a "temporal brush" with which an analyst could highlight a section of a timeline (e.g. a two day sub-period), with the result being the display (on linked scatterplots and map) of all data representing the time span selected. Brushing, as implemented in EDA applications, typically allows an analyst to highlight specific individual entities, while focusing allows analysts to select a value range within which all included entities will be highlighted. Thus, Monmonier's "brushing" corresponds more closely to the EDA method of "focusing" since it involves highlighting a value range from a continuous vector (i.e. the timeline).

Focusing directed on attribute values rather than time has also been applied to spatiotemporal data. For example, MacEachren et al. (1997) combined attribute focusing with animation so that an analyst could focus on a particular subset of mortality data (e.g. the top 20% of the data range) and observe changes in the spatial location of that subset over time. An assessment of this method indicated that it successfully prompted domain experts to notice space-time patterns they otherwise may have missed (MacEachren et al 1997). Brushing (in the EDA sense of selecting a potentially non-contiguous set of entities that will be highlighted) has been applied to the temporal component of data through manipulable legends that allow analysts to highlight specific times, such as a particular hour or day (Edsall and Peuquet 1996).

Learning and problem solving with maps

Learning and problem solving with maps has been the focus of research by cartographers, psychologists, educators, and others. Issues relevant to the current project include: the cognitive prerequisites for learning and problem solving with maps (e.g. Winn 1987; Liben and Downs 1989; Liben and Yekel in press); memory for mapped information (e.g., Eastman 1985; Gilhooly et al. 1988; Mersey 1990); strategies for successful learning and problem solving (e.g. Thorndyke and Statz 1980; Crampton

1992); and the implications of representational choices for learning and problem solving processes (e.g. Kulhavy and Schwartz 1980; Yarnal and Coulson 1982; Peterson 1985).

While there is a substantial body of research on these topics, limited attention has been given to the potential impact of dynamic maps on learning and problem solving. For non-interactive dynamic maps, MacEachren (1992) demonstrated a map learning strategy using sequential presentation of an organized set of subcomponents made it easier to acquire knowledge about city structure. Slocum and Egbert (1993), in contrast, found that sequencing choropleth maps that were segmented by category neither helped nor hindered map learning, but it did increase response times for tasks based on the learned map. In contrast, for temporal information, Koussoulakou and Kraak (1992) found that an animated display resulted in faster response times than did static maps. It appears that animated maps that are not user controllable result in more complete learning, but that the success of these displays depends on the kind of information to be learned and/or the design of the displays.

Edsall et al. (1997) examined the affect of user interaction on effectiveness of animated maps depicting spatiotemporal data. They focused was on animated maps with manipulable legends that allowed users to quickly jump to different points in time. Performance using two versions of their legend was assessed — one that treated time as linear, the other as cyclic. No significant differences in accuracy of response to simple information retrieval questions were found, but individuals clearly used very different strategies for exploring the information.

THE EARTH SYSTEMS VISUALIZER (ESV)

This discussion of the goals of ESV is organized using the 3-tier structure articulated by Howard and MacEachren (1995), focusing on conceptual goals (what tasks the system was designed to facilitate), operational goals (how these goals were operationalised with specific tools), and implementation strategy (the “look and feel” of the system).

Conceptual Level Goals

ESV was designed with two primary conceptual level goals: to facilitate science learning and to examine the cognitive processes that underlie science learning. In relation to the first, the key goal was to prompt student understanding of relationships among space, time, and attributes. Our primary objective was to develop a tool that allows students to recognize the distinction between linear and cyclic weather processes. A secondary objective was to allow users to explore the relationships among multiple geographic phenomena.

Operational Level Goals

Data in ESV are stored as semi-transparent raster layers that users can turn on and off to create visual overlays (**Figure 2**). Because raster-based data layers are traditionally opaque, they normally are viewed separately, or, when commensurate, as an average or composite. ESV, in contrast, uses semi-transparent data layers that are designed to facilitate learning about relationships among phenomena, both spatially and temporally. For example, ESV can be used by students to explore the spatial association between clouds and surface air temperature. Moreover, because ESV supports spatiotemporal data, students can look for a possible lag period in the relationship between clouds and air temperature.

In a dynamic cartographic representation, temporal legends can serve a dual role: to decode the symbols used to represent phenomena and as an interactive control for the display. Earth science processes exhibit both long-term or linear trends (e.g. global warming) and recurring or cyclic patterns (e.g. diurnal temperature fluctuations or seasonal variations in precipitation). ESV addresses each of these by using a *linear* legend to denote days of the year and a *cyclic* legend to denote times of day (**Figure 1**). The linear legend allows the user to focus on a subset of time by adjusting the start and end dates of an animation segment (**Figure 3**). As implemented here, the temporal delimiters used for focusing apply the metaphor of indent tabs on a typewriter or a word processor.

A separate control linked to the cyclic legend allows the user to brush on particular times of day (**Figure 4**).

Extending the EDA concept of brushing into the temporal domain allows users to search for geographic *patterns* that may appear only at certain times. For example, if one is interested in changes in daily maximum temperatures over a one-week period, the ability to suppress the dominant (and potentially overwhelming) diurnal temperature cycle is very useful. As a result, more subtle patterns or longer-term trends might visually emerge from "the noise." As visualization environments become increasingly data-rich, the ability to filter data using brushing becomes essential.

Implementation

Three data layers are included in the ESV prototype: land temperature, sea-surface temperature, and cloud cover. Other data layers could be incorporated easily into the system. The data included have a temporal resolution of six hours, a spatial resolution of approximately 50 kilometers, and a temporal extent of one week (February 10-16, 1998). ESV was built using Macromedia Director (version 6.5), a popular commercial multimedia development tool. ESV was designed to perform well with novice users working with low-end computer displays. The display size was set to 640 x 480 pixels, and all the graphic elements were set to 8-bit color (256 colors). By utilizing Macromedia Shockwave technology, students can use ESV without special software (it plays in their web browser). ESV can downloaded at http://www.geovista.psu.edu/grants/VisEarth/animations1.html.

ASSESSMENT USING FOCUS GROUPS

Our objective in initial testing was to evaluate and refine ESV in preparation for its use in formal cognitive testing. Cognitive cartographic research often fails because map/system design flaws are not identified until expensive and time-consuming user testing has already been completed. A focus group methodology was employed in order to minimize the chance that unintended variables or distracting interface flaws would prevent us from answering our research questions.

Focus Groups: An Overview

Focus groups, sometimes called structured discussions, do not measure effectiveness and are not a tool for hypothesis testing. Instead, they are a research methodology for assessing a product or service by soliciting qualitative feedback from small groups (five to twelve participants) in an informal yet structured environment (Morgan and Krueger 1993; Krueger 1994; Monmonier and Gluck 1994). The use of a discussion facilitator and a preplanned "protocol" to direct the group discussion maximizes participant feedback and keeps the group directed to the topics at hand. Open-ended questions, called *probes*, are used to stimulate discussion and focus on important issues. Examples include "Can you be more specific?" and "How could this be done differently?" The point of a focus group is not to reach consensus, or to solve a specific problem, but rather to understand how and why people respond to something, focusing on both the specific (i.e. the use of a certain color) and the comprehensive (i.e. participants' holistic impressions of the system).

Focus groups are a cost-effective way to generate qualitative evidence concerning the pros and cons of a map or geovisualization system. This method requires less time than other forms of qualitative assessment because data collection is done in groups. Additionally, the open-ended nature of the questions allows for unexpected results to emerge, which is a major advantage over traditional closed-form surveys.

Focus Group Sessions

Focus groups must be small enough to allow everyone the opportunity to speak, yet big enough that diverse opinions can be heard. Also, participants should also regard each other as equals to prevent certain members from feeling intimidated or allowing others to dominate. For these reasons, we did not mix expert and novice users in the same sessions. Instead, there were ran two separate sessions. One group of six included faculty and senior graduate students from the Penn State Department of Geography, all with considerable research and teaching expertise in cartography and visualization. A second group of eight included undergraduate students who had little or no experience with cartography or geographic visualization.

Session Format

Each session lasted approximately one hour and was organized into three sections: (i) a five-minute verbal introduction and demonstration of ESV by the facilitator; (ii) a ten-minute period during which participants were asked to use ESV (each participant had their own computer); and finally, (iii) the group interview. Although the facilitator took notes, the session was also recorded on tape for later review. The use of pre-scripted grouped questions helped to ensure that all important topics were covered in the time allotted, although the protocol was flexible enough that productive digressions could be accommodated. In general, each question generated five to ten minutes of discussion, during which each participant was asked to voice an opinion. The results section below contains specific examples of questions and responses.

The Facilitator

Two of the authors (one the system designer) served as focus group facilitators. Traditional focus group rationale contends that participants will be more forthcoming and critical if the designers or authors are not present. However, more recent thinking (Morgan and Krueger 1993) counters this notion. First, the product designer has intimate knowledge of the product. Second, he or she can converse fluently with any "expert" participants involved. Third, the designer can steer the conversation away from ideas that might be interesting but impossible to implement. It is important to remember that the focus group is not a “brainstorming session” but rather a carefully planned discussion designed to improve the product. Although the participants knew the facilitators, they were encouraged to be as critical and outspoken as possible and had little difficulty in formulating and expressing their opinions.

Results from the Focus Groups

An important difference between focus groups and most other assessment methods is that participants interact with each other. Listening to others express their ideas helps participants formulate their own. Indeed, a “trigger effect” was witnessed when something one participant said generated a flurry of responses from other participants. At times, participants disagreed and would talk about the issues amongst themselves with little or no prompting from the facilitator. After an hour, several key points about the interface and the temporal legends emerged from the sessions. These are summarized below.

Evaluating the System: The Interface

When asked “Did anyone get confused by the interface?” the answer was a unanimous no. This was followed up with the probe “Was it easy to tell if something was ‘on’, or currently selected?” The feeling among the group was that the consistent use of red to indicate "on," and white to indicate "off" was intuitive. One participant noted that ESV used three iconic styles for depicting "on / off" ("radio button", "check box", and "menu selection") and was concerned that this might cause confusion. The facilitator followed up on this with the question *are the different ways of depicting on/off a problem?* to which the group responded either that they had not noticed or did not think it was a problem because each of the three on / off icons are commonplace. The issue was further investigated by asking “Did anything about the interface behave differently than you expected?” to which most of the participants said that temporal brushing was inconsistent with the behavior of other controls (when activated, the animation would stop and ask for user input). Furthermore, they did not like instruction screens that automatically appeared (such as when temporal brushing was activated) and asked for a less intrusive alternative. There was a universal request for better navigation controls, including a step-by-step freeze-frame VCR style button and a pause button. Overall, participants had a positive reaction to ESV and found it easy to use. Earlier versions of ESV showing changes implemented as a result of the focus group suggestions can be seen at (http://www.geovista.psu.edu/grants/VisEarth/olderESV.html).

Evaluating the System: The Temporal Legends

Although the linear temporal legend was understood by the participants, the cyclic legend, however, generated considerable discussion. Our initial cyclic temporal legend was visually distracting and disliked (**Figure 5**). Comments from participants included: “It was the first thing I noticed and I couldn’t ignore it even when I wanted to,” “I never used it. It was very annoying,” and “It is too far (and too small) from the images to be helpful.” Although the cyclic temporal legend was one of the smallest graphical elements on the screen, it was difficult for the eye to ignore because of its strong visual contrast from frame-to-frame combined with continuous apparent rotation. Given that visual perception is especially attuned to movement, any graphic elements that move or rotate will be difficult for the user to ignore (Mowafy et al. 1990). This effect was heightened as the pace of the animation increased, causing the legend to spin even faster.

Participants had other problems with the cyclic temporal legend. Most importantly, they were not sure which part of the globe the legend time referred to. Some participants (novices and experts alike) realized it was not 6 am everywhere at once. Some of these participants, however, assumed that we had standardized (i.e. mosaiced) the images so that it was (which we had not done). This confusion underscores the difficulties in designing a system for communicating global, high-resolution temporal information using relative (i.e. local) measures of time. Although the cyclic legend was disliked, when asked “Should the cyclic legend be removed?” most participants responded “No.” To explore this seeming contradiction, we asked “How could the cyclic temporal legend be improved?” Ideas ranged from moving it closer to the images to combining both the linear and cyclic components into a single hybrid temporal legend. Although available resources did not allow us to address all of the problems identified, we believe that those relevant to subsequent formal testing were dealt with adequately.

Comparing Groups: The Nintendo Generation

After conducting the two focus group sessions, we were surprised by the similarity in reactions, insights, and suggestions made by the expert and novice groups. Although experts were better able to express their concerns, at times citing relevant literature or drawing on personal experience with map and interface design, the basic insights of both groups were the same. In fact, some of the ideas from the novice group sessions showed remarkable sophistication and experience with computer interfaces. Our novice participants seem to fit in the category that Cartwright (1999) calls "the Nintendo generation map user." Although novice users had little or no cartographic/geovisualization expertise, they were generally experienced computer users, and in particular, computer game players. As a result, they learned the interface quickly, felt confident exploring and pushing the interface to its limits, and in general had high expectations of ESV. One important difference, however, did emerge between these groups; when asked to evaluate a sub-component of the interface, novice users responded in the first person (e.g. "I found this confusing") whereas experts thought more broadly of potential users (e.g. "Students might find this confusing").

STRUCTURED USER TESTING

Once ESV was refined, we used structured user-testing to provide insight on (1) whether novice-level users can understand and successfully use temporal brushing and focusing interface controls and (2) the impact of temporal brushing and focusing tools on the student's ability to develop a conceptual understanding of earth-climate processes. More specifically, our methodology was designed to help answer the following questions:

1. What preconceptions do students have about the climate variables presented by ESV and their relationships?
2. Do students understand the interface tools (especially temporal brushing and focusing) and how to apply them for solving tasks?
3. After using ESV, can students generate reasonable hypotheses about climate relationships and how do these hypotheses compare to their pre-conceptions?
4. What affect do temporal brushing and focusing capabilities have on the hypotheses generated?

5. How does prior knowledge about climate influence success in using ESV or the extent to which ESV helps in hypothesis generation?
6. What are the overall student impressions of ESV?

Test Stimuli

Two versions of ESV were created in order to test the potential added value of temporal brushing and focusing over the more traditional VCR-style control of animations, one with and one without brushing and focusing tools. Below, the one with these tools is referred to as the "enhanced" version (see Figure 1).

Participants

Because ESV was designed as an educational tool for entry-level college earth science students, we tested undergraduate students as participants whose age and experience most closely matched our target audience. Our subjects included thirty-four undergraduate students with equal numbers of male and female students, average age of 20.6 years, and an average of two years of college. Seventeen subjects were assigned randomly to the standard and enhanced ESV groups. Within the standard ESV group a few more subjects were earth science majors (11 versus 9) while the mean number of courses focused on climate was slightly lower (0.76 versus 1.1).

Questions

A set of questions was devised to determine the extent to which each test group could retrieve information from the maps, identify patterns they saw in the maps, and generate hypotheses about the processes driving the patterns they observed in the maps.

Open-ended short-answer questions (**Table 1, questions 1 and 2**) were used to assess each student's understanding of climate variable relationships both *before* and *after* the ESV session. This allowed us to relate performance with ESV to preconceptions about climate and to determine whether focusing and brushing tools had an impact on this relationship.

In order to establish a baseline of understanding for each student, an open-ended question about the general relationship between clouds and temperature was asked *before* the students saw ESV (**Table 1, question 1**). Two related, more focused questions were asked *after* the ESV session. Although these questions, necessarily, lead to responses that were more specific than those of the pre-ESV question, they allowed us to examine how the sophistication of responses (relative to the group as a whole) prior to using ESV compared to that after using ESV. They allowed us to judge whether any differences in these answers were related to use of brushing and focusing tools.

Users of the enhanced ESV were also asked two additional questions in which they were prompted to explain why brushing and focusing might be used. Answers to these questions helped us assess whether students understood the purpose of the tools.

Table 1. Open-ended short answer questions

Questions asked of all participants
1. How are clouds and temperature related? (Pre-test question) 2. In two or three sentences, state a hypothesis about the relationship between clouds and temperature during this week. (Post-ESV question) 3. Describe how the land temperatures change throughout the week. (Post-ESV question)
Additional questions asked only of participants that used the enhanced ESV
4. I think temporal focusing is most useful for … 5. I think temporal brushing is most useful for …

Six multiple choice questions were used to determine whether students could interpret basic elements of ESV. These questions required students to retrieve information and interpret patterns from the mapped data (e.g. *"In which direction do cloud formations pass over South America?"*). Along with the short-answer questions described below, these questions helped to characterize each subject's general level of knowledge about climate phenomena; since ESV should help enable students to answer questions about particular places, specific points in time, and attributes of processes represented. The six information retrieval and pattern identification questions were designed specifically to address these three components systematically. Pairs of questions focused on time, space or attributes, while holding the other two constant. For example, *Which continent is* *hottest in the morning on February 14?* forced students to compare the same attribute (temperature) at different places for the same time period.

The multiple choice questions for the standard and enhanced ESV groups were identical, except for three additional questions posed to enhanced ESV users related to brushing and focusing. These questions tested whether students understood how to use the temporal brushing and focusing tools (e.g. *"If I want to see only the temperatures at 6 a.m., I should …"*). The complete set of questions can be viewed online at http://www.geovista.psu.edu/grants/VisEarth/survey.html

A third component of testing focused on obtaining subjective responses to the system as a whole and to its components. To do this, we created 21 semantic differential word pairs that could be used to describe the subject's overall reaction to three aspects of ESV: subjects were asked to rate the controls, the map and the data by choosing a position along a number line from 1 to 7. One word from each pair was placed at the end of each number line (e.g. very fast = 1, very slow = 7). The full set of word pairs is included with results in **Table 5**. For previous cartographic application of this technique see Harrower et al. (1997), Gilmartin (1978), and Petechnik (1974). To avoid leading the participant, the positive-negative polarity of the word-pairs was randomly assigned.

Experimental Session

Two of the authors acted as facilitators for each of the four testing sessions. The sessions were held in one of our departmental computer labs, with twelve Pentium computers, set up in rows of four. Seven to ten subjects participated at a time, but worked on their own. Subjects were randomly assigned to use either the standard or enhanced ESV. Students were given a short explanation of the study and a brief overview of the testing procedure.

Participants began by generating an initial response to the question: *How are clouds and temperature related?* Their responses were collected before subjects saw ESV. Subjects were then given a second sheet containing the semantic differential questions and a short biographical questionnaire. To minimize and standardize the interaction between the participants and the facilitators, subjects were introduced to ESV through a series of on-screen instructions. Participants were told there was no time limit for the session and were encouraged to spend as much time as needed to familiarize themselves with the system before answering any questions.

Data Collected

Once the subjects were ready, they proceeded with the session. Questions appeared at the bottom of each screen prompted participants to choose an answer (for multiple choice questions) or enter a text response (for open-ended questions) and responses were recorded automatically. Also recorded were the time elapsed between answers and the sequence of tools that the participant used between answering one question and the next. Due to technical constraints, however, tool use was only recorded while students answered multiple-choice questions. At the end of the session, subjects were asked to rate the system and its components using the semantic differential word pairs and to complete the biographical form.

Analysis and Results

Open-ended Questions

To analyze responses to the open-ended questions, we developed three criteria for consistently judging the characteristics of written responses:

1. ***Consistency*:** Was the hypothesis or description consistent (logical/testable)? The hypothesis does not necessarily have to be correct, rather it should be testable and logically consistent. ESV is designed to prompt students to investigate climate phenomena, rather than always produce a correct answer. It is, thus, an exploratory tool.
2. ***Dimensionality*:** Did the participant refer to space, time and process in their answer? The goal of ESV is to facilitate analysis of space, time and attributes together. We were interested in whether participants initial- and post-ESV responses included reference to each.

3. ***Confidence***: What was the participant's level of confidence in answering the question? For example, answers beginning with "I'm not sure, but …" or finishing with a question mark were judged to be less confident.

Responses to each of the open-ended questions were rated on these three criteria, using a 4-point scale for *consistency*, 3-point scale for *dimensonality* and a 2-point scale for *confidence* (1=confident). Incorrect and nonsensical participant responses, and those not containing a hypothesis, were assigned a consistency rating of zero; those that were logical/testable but very simplistic were given a rating of 1; those that were logical/testable, but not very comprehensive were given a rating of 2; and those that were both logical/testable and comprehensive were assigned a rating of 3. When participant answers were rated on *dimensionality*, if the answer made reference to space, time and process, it received a rating of 3. An answer that referred to only two of these qualities received a rating of 2, and those that made reference to only one of these characteristics were given a rating of 1. Examples of student answers to specific questions are presented in Table 2.

Table 2. Examples of student answers and their rating.

Consistency	***How are clouds and air temperature related?***
Invalid	"I don't know"
Novice (score 1)	"More clouds when cold than warm?"
Intermediate (score 2)	"Clouds reflect incoming solar radiation. They also keep warm air between the earth and the clouds."
Advanced (score 3)	"Clouds can act to either increase or decrease the earth's temperature. Clouds can either reflect the sun's rays, keeping radiation from reaching the earth, or they can absorb the radiation coming off the earth's surface and keep the air below the clouds warmer. Clouds also carry fronts of weather."
Dimensionality	***In two or three sentences, state a hypothesis about the relationship between clouds and temperature during this week.***
Novice (score 1)	"The more clouds over the land, the hotter the temperature. This may be due to the greenhouse effect."
Intermediate (score 2)	"Generally where there are more clouds, the land temperature stayed warmer. Clouds do not affect the water temperature as much as they do the land temperature."

Advanced (score 3)	"The cloud cover in the north reflected enough radiation so that the sun could not warm up the land anymore. The small amount of cloud cover in the equatorial region had the effect of cold nights and warm days. This is because the radiation was not able to reach the earth when the sun shone during the day and at night the clouds were not there to trap the warmth in."
Confidence	***In two or three sentences, state a hypothesis about the relationship between clouds and temperature during this week.***
Not Confident	"I think the cloud cover in general makes the temperature lower. I really don't see a pattern well enough to say anymore."
Confident	"Warmer in the eastern hemisphere later in the week."

As mentioned above, one of the goals of ESV was to prompt greater awareness of geographic dimensions of space, time and attribute. To test for this, a series of Wilcoxon matched-pair signed-rank tests compared pre-ESV and post-ESV written responses. A statistical difference at the p = .01 level was found for the *dimensionality* criteria when the answers to both question 2 and question 3 (post-ESV questions) were compared to the answer to question 1 (pre-ESV question). When comparing question one (pre-ESV) to question 2 (post-ESV), almost half of the participants (eight of the seventeen) generated more dimensionally complex answers after using ESV. A preliminary conclusion that can be drawn, then, is ESV was able to prompt students to think to a greater degree about aspects of time, space and attribute then they did without ESV (i.e. before the session). There were no statistically significant differences between the *confidence* or *consistency* of the pre-ESV answer and the post-ESV answers.

Table 3. Wilcoxon matched-pairs signed-rank test results for comparison of answers to pre- and post ESV questions.

	Confidence		**Dimensionality**		**Consistency**	
Question	**Wilcoxon Statistic**	**Signif.**	**Wilcoxon Statistic**	**Signif.**	**Wilcoxon Statistic**	**Signif.**
1 vs. 2	-1.000	0.3173	-3.9117	0.0001**	-1.6063	0.1082
1 vs. 3	-.3333	0.7389	-3.7503	0.0002**	-1.9571	0.0503

* indicates statistically significant result (at p = 0.05)
** indicates statistically significant result (at p = 0.01)

As indicated above, subjects who used the enhanced ESV were asked two additional open-ended short-answer questions to assess their understanding of the temporal focusing and brushing controls (see Questions 4 and 5, **Table 1**). Six of the seventeen participants (35.3%) were unable to describe a reasonable potential use of temporal focusing. In contrast, only one student (5.8%) was unable to describe a correct potential use of temporal brushing.

There were no significant differences in the written responses between the standard and enhanced ESV groups (looking at consistency, dimensionality or confidence). Thus, we can conclude that the enhanced ESV was not as successful at promoting hypothesis generation as we had hopedThis may be because many students did not recognize the need to or did not understand how to use the enhanced ESV to select and filter out information that distracts from overall trends or relationships between variables (e.g. diurnal temperature fluctuations when looking for a weekly temperature trend). This confusion in how to use ESV effectively could have been addressed by better instruction screens, especially with temporal focusing (as 35.3% were judged to not understand how to use the tool).

In response to the question "*Describe how the land temperatures change throughout the week.*" many subjects made no reference to the weekly trend at all, fixating instead on the strong diurnal pattern. This pattern of fixating on the strong diurnal temperature pattern might be expected in those subjects with no access to brushing (those using the standard ESV), as they had no filtering mechanism with which to eliminate the overwhelming diurnal fluctuation. Those students who did have access to temporal brushing could have used it to suppress these fluctuations, and were expected to give higher-quality answers to this question. Interestingly, when asked direct questions about the purpose and/or utility of temporal brushing, subjects generally gave satisfactory answers. However, understanding the purpose of a tool and recognizing when it is useful for solving a problem are different.

Analysis and Results

Multiple Choice Questions

The multiple choice section of our test included six questions designed to determine whether students could retrieve basic facts from—and notice and interpret simple patterns in—the animated maps. All participants were grouped on two variables: (1) according to whether they used the standard or enhanced ESV (enhanced ESV, n = 17; standard ESV, n = 17) and (2) whether they had taken any classes that focused on climate or weather (at least one class, n = 18; no classes, n = 16).

The 17 participants who used the enhanced ESV were grouped on two additional characteristics, both related to confusion about tool use. A participant was judged to be confused if (a) they could not give a reasonable example of a potential use for each tool in the written questions (confused, n = 6; not confused, n = 11), or (b) they did not answer all three multiple choice questions that dealt with the tools correctly (confused, n = 5; not confused, n = 12).

We conduct two tests for differences among groups. A Mann-Whitney test (a non-parametric version of the t-test) was used to test for differences between two groups (**Table 4**). Those participants who had taken relevant classes exhibited significantly better overall performance on the multiple choice questions (measured as the mean number correct for multiple choice questions 1-6). No significant differences in mean percent correct for the multiple choice questions were found between either the enhanced and standard ESV groups or between the confused and non-confused groups. A significant difference in performance was found between the enhanced and non-enhanced ESV group for question 2 (*"On which day is Africa the hottest?"*) at the $p=0.05$ level, but because we used multiple statistical tests, $p=0.01$ is more appropriate according to the Bonferroni correction (Dai and Van der Maarel 1997).

Table 4. Mann-Whitney results for correct answers of multiple choice questions

Grouping	Question 1		Question 2		Question 3		Question 4	
	Z	**P**	**Z**	**P**	**Z**	**p**	**Z**	**p**
Enhanced vs. Standard	-1.000	.3173	-2.061	.0394*	-.3830	.7017	0.000	1.000
Classes vs. No Classes	-1.060	.2888	-1.092	.2747	-.1805	.8567	-1.784	.0745

Grouping	Question 5		Question 6		All Questions	
	Z	**p**	**Z**	**p**	**Z**	**p**
Enhanced vs. Standard	0.000	1.000	-.6770	.4980	-0.84	.3790
Classes vs. No Classes	-1.704	.0883	-1.037	.2996	-2.01	.0439*

* indicates statistically significant result (at p = 0.05)
** indicates statistically significant result (at p = 0.01)

Our expectation prior to testing was that students who had access to brushing and focusing tools would perform better in the multiple-choice questions. Regarding the overall number of questions answered correctly, however, no statistically significant differences were observed between the groups using the standard and enhanced ESV (**Table 4**, 'All Questions'). The average number of correct answers for the enhanced group was 4.23/6.00 (70.5% correct) versus 4.06/6.00 for the standard group (67.7% correct). This may be explained by noting that there was a clear bimodal distribution in the enhanced group responses. Those students who were confused by temporal focusing did poorly in the skill testing questions (60% correct). In contrast, students who understood and used the tools generated some of the highest test scores (75% correct). Four of the five students who obtained perfect scores used the enhanced ESV. Thus, it appears that the enhanced ESV may assist understanding of spatiotemporal phenomena, but require training for some students before this benefit can occur.

Analysis and Results

Semantic Differential Word Pairs

Twenty-one bi-polar word pairs were used to obtain participant ratings, on a 7-point scale, of (a) the controls, (b) the map and (c) the phenomena depicted in the visualization (**Table 5**). Using a Kruskal-Wallis test for differences between groups, no significant differences in how participants characterized the interface, map or phenomena were found between users with the enhanced or standard ESV. Similarily, the number of climate or weather classes attended produced no significant differences.

Significant differences in users' reactions to ESV were found when performance on multiple choice questions was used as the grouping factor (**Table 5**). One pattern that emerged was that individuals who performed at the medium level (4-5/6 correct) had the most positive reaction to the interface; they thought the interface was more attractive, better organized and less confusing than did either the advanced or novice-level users. The high-performance group (6/6 correct) had a similar, generally positive, reaction to ESV. The low-performance group (3 or less correct), not surprisingly, was unimpressed with the ESV. Why these individuals had trouble answering questions with the ESV remains unanswered. This is clearly one area where future research efforts will have to be directed.

Table 5. Mean ratings and Kruskal-Wallis results of semantic differential word pairs, by level of accuracy in answering multiple-choice questions

Rating the Interface	**Low Accuracy**	**Medium Accuracy**	**High Accuracy**	**Chi-square**	**p**
Fast-slow	2.33	2.82	2.60	1.53	.4654
Unattractive-attractive	3.33	5.82	5.80	11.31	.0035**
Unorganized-organized	3.17	5.82	5.60	12.77	.0017**
Clear-confusing	3.33	2.06	3.00	7.66	.0217*
Helpful-hindering	2.75	1.71	2.40	10.93	.0042**
Easy to understand-Difficult to understand	2.75	1.76	2.20	7.34	.0254*
Rating the Phenomena	**Low Accuracy**	**Medium Accuracy**	**High Accuracy**	**Chi-square**	**P**
Hot-cold	3.00	3.41	2.80	2.02	.3640
South-north	3.18	3.94	3.20	3.58	.1663
West-east	4.00	3.47	3.40	0.66	.7174
Slow-fast	4.00	5.12	5.40	7.61	.0222*
Regular-random	2.81	2.82	1.40	7.05	.0293*
Inactive-active	4.42	5.71	5.40	6.45	.0398*
Stormy-clear	4.08	4.53	4.80	1.00	.6062
Rating the map	**Low Accuracy**	**Medium Accuracy**	**High Accuracy**	**Chi-square**	**P**
Misleading-truthful	4.33	5.76	6.20	6.04	.0489*
Simple-complex	3.64	2.76	4.60	4.46	.1070
Incomplete-complete	4.00	5.29	5.00	3.40	.1823

Specific-generalized	3.75	4.41	4.40	1.44	.4864
Worthless-valuable	4.75	5.82	5.20	6.57	.0374*
Clear-vague	3.42	2.06	2.60	6.60	.0368*
Easy to understand – difficult to understand	3.33	2.00	2.60	5.40	.0670
Meaningful-meaningless	2.58	1.82	2.20	4.40	.1104

* indicates statistically significant result (at p = 0.05)

** indicates statistically significant result (at p = 0.01)

We also wanted to know how students rated their own skills (i.e. their *confidence*) in using the ESV. In order to determine this, participants were asked to rate their ability to use the ESV on a scale from one to four. Using a Mann-Whitney test, we found no significant relationship between self-reported level of confidence and how well participants did on any components of the test.

CONCLUSIONS

Two temporal legend styles were incorporated into ESV in order to promote an understanding of time as both linear and cyclic. The linear temporal legend was favorably reviewed by the focus group participants, but, the cyclic temporal legend was generally disliked. Two specific problems emerged from the focus group session. A rapidly spinning graphic element was visually distracting. This component was redesigned to make changes from timeframe to timeframe much less dramatic. The second problem is more difficult and without such an easy solution; although local time (i.e. 3 p.m.) is a familiar and easily understood unit of measurement, the focus group revealed that using such relative measurements can cause confusion and may not be appropriate for *global* data sets. Thus far, we have developed no solution to this problem.

Overall, focus group participants liked using the ESV and felt that our design goals had been implemented successfully. The participants were, however, able to identify inconsistencies in the design of these tools, specifically the behavior of the system upon activation of the tools, and thus requested more flexible speed controls, new VCR buttons including frame-by-frame advance, the elimination of some instruction screens, and an exit button.

One important difference between the focus group sessions and the formal testing concerned the temporal focusing control. In the formal tests, slightly more than one-quarter of the participants who had the enhanced version of ESV did not seem to understand how to use temporal focusing, or what it should be used for. This was a much higher percentage than we expected after the positive reaction that both the novice and expert participants had to temporal focusing in the focus groups. The difference between these two sessions, apparently, was the ability of users in the focus group to interact with the instructor and ask questions. In other words, although some of the focus group participants may have been confused initially about the temporal focusing controls, a short verbal instruction or demonstration ameliorated this confusion. Given that ESV is designed as a teaching aid in the classroom, the instructor could address such confusion.

A central insight of this research is that positive reactions in focus groups do not (necessarily) translate into good performance with an interface, and novel interfaces may not result in improved performance unless sufficient training is provided in how to use them.

As a whole, our results suggest that the level of a visualization system must be well matched to the level of the user. Students who were able to generate reasonable observations about earth-climate processes commented that the system was not especially helpful. Novice-level users also rated the system poorly. The intermediate-level students had the most positive (qualitative) reaction to the system. Across groups, students who did well with information retrieval and pattern interpretation also responded to the system favorably.

Our results also indicate that users presented with tools they do not fully understand are likely to perform more poorly than when not provided with any special tools. Students who were confused by the temporal focusing controls performed poorly on information retrieval and pattern interpretation questions and characterized both the controls and the map negatively. Not only is it necessary that students understand the purpose of a tool, but also that they recognize situations for applying tools. An important component of instruction in tools use, then, may be illustrating their appropriate application via a set of examples.

ACKNOWLEDGEMENTS

We would like to acknowledge support for this research provided through The Visualizing Earth Project (NSF grant #RED-9554504). We would also like to thank Jeff Balmat and Milissa Orzolek for volunteering to participate in the pilot study.

REFERENCES

Acevedo, W. and P. Masuoka (1997). Time-series animation techniques for visualizing urban growth. *Computers & Geosciences* 23(4): 423-436.

Becker, R. A., W. S. Cleveland, and A. R. Wilks (1988). Dynamic graphics for data analysis. In W. S. Cleveland and M. E. McGill (Eds.). *Dynamic graphics for statistics*. Wadsworth & Brooks, Belmont, California.

Buttenfield, B. P. and C. R. Weber (1994). Proactive graphics for exploratory visualization of biogeographical data. *Cartographic Perspectives* 19: 8-18.

Campbell, C. S. and S. L. Egbert (1990). Animated cartography: Thirty years of scratching the surface. *Cartographica* 27(2): 24-46.

Carr, D. B., R. J. Littlefield, W. L. Nicholson, and J. S. Littlefield (1987). Scatterplot matrix techniques for large databases. *N. Journal of the American Statistical Association* 82(398): 424-436.

Carter, R. J. (1996). Television weather broadcasts: Animated cartography aplenty. *Proceedings of the Seminar on Teaching Animated Cartography*, Escuela Universitaria de Ingeniera Tecnica Topografica, Madrid, Spain, August 30 - September 1,1995. http://nvkserver.frw.ruu.nl/html/nvk/ica/madrid/carter.html.

Cartwright, W (1999). The Development of a Hybrid Discrete/Distributed Interactive Multimedia Package for Teaching Geographical Concepts by Exploration. Proceedings of the *19th International Cartographic Conference*, Aug. 14-21, 1999, Ottawa.

Cook, D., J. J. Majure, J. Symanzik, and N. Cressie (1996). Dynamic graphics in a GIS: Exploring and analyzing multivariate spatial data using linked software. *Computational Statistics* 11: 467-480.

Crampton, J. (1992). A cognitive analysis of wayfinding expertise. *Cartographica* 29(3/4): 46-65.

Dai, X. and E. Van der Maarel (1997). Transect-based patch size frequency analysis. *Journal of Vegetation Science* 8 (6): 865-872.

Dorling, D. and S. Openshaw (1992). Using computer animation to visualize space-time patterns. *Environment and Planning B: Planning and Design* 19: 639-650.

Eastman, J. R. (1985). Graphic organization and memory structures for map learning. *American Cartographer* 22: 1-20.

Edsall, R. M., M.-J. Kraak, A. M. MacEachren, and D. J. Peuquet (1997). Assessing the effectiveness of temporal legends in environmental visualization. *GIS/LIS '97 Cincinnati*, OCT. 28-30, pp. 677-685.

Edsall, R. M. and D. J. Peuquet (1996). Graphical query techniques for temporal GIS. *Proceedings: ACSM/ASPRS Annual Conference Seattle*. WA, pp. 182-189.

Gersmehl, P. J. (1990). Choosing tools: Nine metaphors for four-dimensional cartography. Cartographic Perspectives, Spring: 3-17.

Gilhooly, K. J., M. Wood, P. R. Kinnear, and C. Green (1988). Skill in map reading and memory for maps. *The Quarterly Journal of Experimental Psychology* 40A: 87-107.

Gilmartin, P. P. (1978). Evaluation of thematic maps using the semantic differential test. *American Cartographer* 5(2): 133-139.

Gould, P., D. DiBiase, and J. Kabel (1990). Le SIDA: la carte animee comme rhetorique cartographique appliquee. *Mappe Monde* 90(1): 21-26.

Hägerstrand, T. (1974). Tidsgeografisk beskrivning. Syfte och postulat (Time-geographical descriptions. Purpose and postulate).

Harrower, M., C. P. Keller, and D. Hocking (1997). Cartography on the Internet: thoughts and a preliminary user survey. *Cartographic Perspectives* 26: 27-37.

Howard, D. and A. M. MacEachren (1995). Constructing and evaluating an interface for visualizing reliability. *Proceedings of the 17th International Cartographic Conference*, Barcelona, Spain, September 3-9: 320-329.

Koussoulakou, A. and M.-J. Kraak (1992). Spatiotemporal maps and cartographic communication. *Cartographic Journal* 29(2): 101-108.

Kraak, M.-J. and A. M. MacEachren (1994). Visualization of spatial data's temporal component. Proceedings, *Spatial Data Handling, Advances in GIS Research*, Edinburgh, Scotland, 5-9, September, 1994, pp. 391-409.

Krueger, R. A. (1994). *Focus Groups: A practical Guide for Applied Research* (2nd Ed.). Sage Publications, Thousands Oaks, California, 1-255pp.

Kulhavy, R. W. and N. H. Schwartz (1980). Mimeticism and the spatial context of a map. *Bulletin of the Psychonomic Society* 15: 416-418.

Liben, L. S. and R. M. Downs (1989). Understanding maps as symbols: The development of map concepts in children. *Advances in Child Development and Behavior* 22: 145-201.

Liben, L. S. and C. A. Yekel (in press). Preschooler's understanding of plan and oblique maps: The role of geometric and representational correspondence. *Child Development*.

MacEachren, A. M. (1992). Application of environmental learning theory to spatial knowledge acquisition from maps. *Annals of the Association of American Geographers* 82(2): 245-274.

MacEachren, A. M. (1995). *How Maps Work: Representation, Visualization and Design*. New York: Guilford Press.

MacEachren, A. M. and D. W. DiBiase (1991). Animated maps of aggregate data: Conceptual and practical problems. *Cartography and Geographic Information Systems* 18(4): 221 229.

MacEachren, A. M., D. Howard, M. von Wyss, D. Askov, and T. Taormino (1993). Visualizing the health of Chesapeake Bay: An uncertain endeavor. Proceedings, *GIS/LIS '93*, Minneapolis, MN, 2-4 Nov., 1993, pp. 449-458.

MacEachren, A. M., C. Polsky, D. Haug, D. Brown, F. Boscoe, J. Beedasy, L. Pickle, and M. Marrara (1997). Visualizing spatial relationships among health, environmental, and demographic statistics: interface design issues. *18th International Cartographic Conference* Stockholm, June 23-27, pp. 880-887.

MacDougall, E. B. (1992). Exploratory analysis, dynamic statistical visualization and Geographic Information Systems. *Cartography and Geographic Information Science*, 19(4): 237-246.

McGuinness, C. (1994). Expert / novice use of visualization tools. In *Visualization in Modern Cartography*, ed. A. M. MacEachren and D. R. F. Taylor, pp. 185-200. London: Pergamon.

Mersey, J. E. (1990). Colour and Thematic Map Design: The Role of Colour Scheme and Map Complexity in Choropleth Map Communication. Monograph 41, *Cartographica* 27(3).

Moellering, H. (1976). The potential uses of a computer animated film in the analysis of geographical patterns of traffic crashes. *Accident Analysis & Prevention* 8: 215-227.

Monmonier, M. (1989). Geographic brushing: Enhancing exploratory analysis of the scatterplot matrix. *Geographical Analysis* 21(1): 81-84.

Monmonier, M. (1990). Graphic scripts for the sequenced visualization of geographic data. Proceedings of *GIS/LIS '89*, 26-30 Nov., 1989, Orlando, FL 381-389.

Monmonier, M. (1990). Strategies for the visualization of geographic time-series data. *Cartographica* 27(1): 30-45.

Monmonier, M. (1992). Authoring Graphics Scripts: Experiences and Principles. *Cartography and Geographic Information Systems* 19(4): 247-260.

Monmonier, M. and M. Gluck (1994). Focus groups for design improvement in dynamic cartography. *Cartography and Geographic Information Systems* 21(1): 37-47.

Morgan, D. L. and R. A. Krueger (1993). When to use focus groups and why, in Morgan, D. L. (ed.) *Successful Focus Groups: Advancing the State of the Art*. Sage Publications, Thousands Oaks, California, pp. 3-19.

Mowafy, L., R. Blake, and J. S. Laapin (1990). Detection and discrimination of coherent motion. *Perception and Psychophysics* 48(6): 583-592.

Petechnik B. B. (1974). A verbal approach to characterizing the look of maps. *The American Cartographer,* 1(1): 63-71.

Peterson, M. P. (1985). Evaluating a map's image. *American Cartographer* 12: 41-55.

Peterson, M. P. (1995). *Interactive and Animated Cartography*. Englewood Cliffs, NJ: Prentice Hall.

Sirken, M., D. Herrmann, and A. A. White (1995). Cognitive aspects of designing statistical maps. In *Cognitive Aspects of Statistical Mapping,* Working Paper Series from the Centers for Disease Control and Prevention, Washington, DC.

Slocum, T. A. (1988). Developing an information system for choropleth maps. Proceedings: *Third International Symposium on Spatial Data Handling,* Sydney, Australia: 293-305.

Slocum, T. A. and S. L. Egbert (1993). Knowledge acquisition from choropleth maps. *Cartography and Geographic Information Systems* 20(2): 83-95.

Szegö, J. (1987). *Human Cartography: Mapping the World of Man*. Stockholm: Swedish Council for Building Research.

Thorndyke, P. W. and C. Statz (1980). Individual differences in procedures for knowledge acquisition from maps. *Cognitive Psychology* 12: 137-175.

Thrower, N. J. W. (1961). Animated cartography in the United States. *International Yearbook of Cartography* 1: 20-30.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46: 234-240.

Winn, W. (1987). Communication, cognition and children's atlases. *Cartographica*, Monograph 36, Atlases for Schools 24(1): 61-81.

Yarnal, C. M. and C. M. R. Coulson, (1982). Recreational map design and map use: an experiment. *Cartographic Journal* 19(1): 16-27.