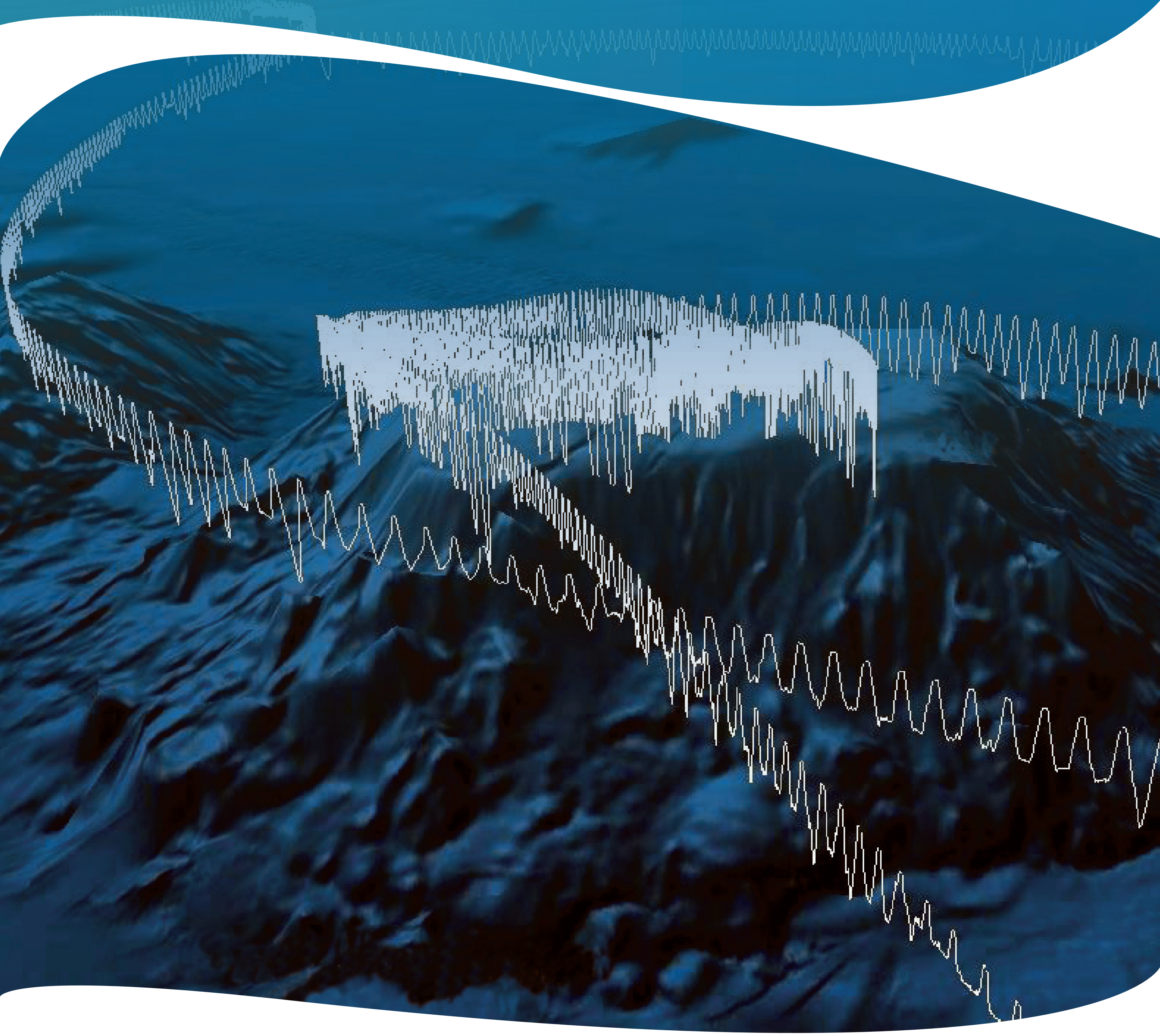
VISUALIZING OCEANS OF DATA
Educational Interface Design
2013 KNOWLEDGE STATUS REPORT



EDUCATION DEVELOPMENT CENTER, INC. • SCRIPPS INSTITUTION OF OCEANOGRAPHY

COVER IMAGE CREATED BY PATRICK ROBINSON

Maxwell, S. M., J.J. Frank, G.A. Breed, P.W. Robinson, S.E. Simmons, D. Crocker, J. Gallo-Reynoso, and D.P. Costa (2012) Benthic foraging on seamounts as a specialized foraging behavior by a deep diving marine mammal. *Marine Mammal Science 28*(3): E333-E344.

REPORT DESIGN

Designed by Jennifer Matthews *(jbmatthews@ucsd.edu)*, Scripps Institution of Oceanography, UC San Diego

Visualizing Oceans Of Data: *Educational Interface Design*

Citation:
Krumhansl, R., Peach, C., Foster, J., Busey, A., and Baker, I. (2012). *Visualizing Oceans of Data: Educational Interface Design.* Waltham, MA: Education Development Center, Inc.

This material is based upon work supported by the National Science Foundation under Grant Nos. 1020002 and 1019644. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

CONTENTS

PREAMBLE

Science is data-intensive, but today's science education is not. In most classrooms, students' work with data is limited to reading graphs prepared by others, or at best collecting simple data sets themselves. While these student-collected data sets allow students to begin building their data proficiency, the conclusions that can be drawn and the lessons that can be learned from these data are limited in scope and can sometimes be compromised by data quality. The large, high-quality scientific data sets that are newly available online allow today's science students to incorporate working with authentic data into their learning experiences, giving them virtually unlimited opportunities to participate in real scientific work.

However, the fact remains that the educational promise of large scientific cyberinfrastructures will not be met without concerted effort. It is a huge leap to bridge from reading graphs or maps that have been carefully prepared to illustrate a particular concept to interpreting data visualizations that may not have ever been seen before, may have data problems, and may not show any obvious trend. It's also a huge leap to bridge from data that students have collected themselves to data that were collected remotely, by instruments students do not understand, in an environment they have not seen.

As one of our advisors, Jim Hammerman (August 22, 2012), noted:

> *It's a really hard and important problem. It shouldn't be so hard for people in schools to use [these professional data sets], but we all know it is. I'm interested in having these sorts of tools available for schools and citizen groups who want to make a difference in the world, making it possible for people to be curious, and making the case for what matters to them using data.*

The Oceans of Data project has made an attempt to define and confront what is "hard" for students and teachers who attempt to use large, online professional data sets. We feel passionately that it's important for us to do this to prepare today's students for tomorrow's world.

THE OCEANS OF DATA PROJECT TEAM

RUTH KRUMHANSL, principal investigator at Education Development Center, Inc.(EDC), provided technical leadership to the project team and coordinated project work with Scripps, the project's advisory board, and NSF. Her focus during implementation of the project was on reviewing and analyzing literature and developing guidelines relevant to Accessing Data and Geo-referenced Data Representations, and she led the synthesis and development of Visualizing Oceans of Data.

CHERYL PEACH, principal investigator at Scripps Institution of Oceanography, played a key role in ensuring the relevance of the Oceans of Data work to scientific cyberinfrastructure projects such as the Oceans Observatories Initiative. In addition, she arranged, co-planned, and hosted the advisory committee meetings, helped to visualize the structure of Visualizing Oceans of Data, and assumed primary responsibility for the dissemination of project findings,

JUNE FOSTER, co-principal investigator at EDC , was instrumental to the conceptualization of the Oceans of Data project, helping to shape the project goals and the research methodologies, and in particular contributing her expertise in Universal Design for Learning to the project. She was primary reviewer of all sections of Visualizing Oceans of Data and lead writer of the Cross-cutting Guideline section Enabling Customization.

AMY BUSEY of EDC was a primary author of Visualizing Oceans of Data. Her particular focus during the literature review and writing was on visual perception and cognitive load theory, and she was lead writer of the related Cross-cutting Guidelines sections. She also researched and wrote the specific guidelines for Animations.

IRENE BAKER of EDC completed the review and coding of literature related to graphs, and was lead author of the Graphs section of Visualizing Oceans of Data. She also conducted interviews with existing Web-based data providers as part of an initial needs assessment.

JACQUELINE DELISI of EDC acted as an internal methodological advisor, advising the project team as they refined the research methodologies, developed coding protocols, and analyzed findings.

KIRA KRUMHANSL of EDC played a critical role by searching for and obtaining literature relevant to the Oceans of Data project work.

ACKNOWLEDGMENTS

We'd like to acknowledge the contributions of our advisors, who shared their considerable experience and insights at two lively and stimulating meetings, as well as in telephone interviews, written comments, and e-mail communications. Their comments on the draft Knowledge Status Report greatly improved its content, particularly in areas where directly relevant literature is sparse. They brought diverse experience in education research, science research, teaching, educational software development, and cyberinfrastructure development to our work, which led to particularly interesting exchanges where we struggled to understand each other's language and perspectives. These productive struggles convinced us that more of these types of conversations are essential if we want to bring expert databases to students.

The Oceans of Data Advisory Board comprised the following members:

YI CHAO, Principal Scientist, Jet Propulsion Laboratory

DANIEL EDELSON, Vice President of Education, National Geographic

ALLISON FUNDIS, Research Scientist and Education and Public Outreach Liaison, Oceans Observatories Initiative RSN, University of Washington

BORIS GOLDOWSKY, Director of Technology, Center for Applied Special Technology

JAMES HAMMERMAN, Senior Researcher and Evaluator, TERC

KIM KASTENS, Doherty Senior Research Scientist, Lamont-Doherty Earth Observatory, Columbia University

JULIANNE MUELLER-NORTHCOTT, Biology and Earth Science Teacher, Souhegan High School

JOHN ORCUTT, Professor of Geophysics, Scripps Institution of Oceanography, UCSD

WILLIAM SANDOVAL, Associate Professor of Psychological Studies in Education, Graduate School of Education and Information Studies, UCLA

We'd also like to thank cognitive scientists Jess Gropen of EDC and Thomas Shipley of the Spatial Intelligence and Learning Center at Temple University for their thoughtful review and insightful feedback that enhanced the quality of this product, the National Science Foundation for funding this work, and our program officer Elizabeth Van der Putten for her support and encouragement along the way.

Introduction

I. INTRODUCTION

I. INTRODUCTION

About the Oceans of Data Project

The practice of science and engineering is being revolutionized by the development of cyberinfrastructures for accessing near real-time and archived observatory data. The NSF-funded project Oceans of Data aims to make it possible for students and their teachers to join that revolution.

The potential exists for classrooms to use state-of-the-art resources and techniques for scientific investigations and to analyze and draw conclusions from many kinds of complex data. But realizing that potential requires breaking new ground. As they stand now, the interfaces and data visualization tools for large science cyber-infrastructure databases are industrial-strength—designed by experts for use by experts—which significantly impedes broad use by novice learners.

What is needed are more "egalitarian" interfaces and data representations that make large scientific databases accessible to, and usable by, nonscientists (some of whom, hopefully, are budding scientists). But doing so is no easy matter for the software developer. Efforts to create interfaces and tools that bridge to the science classroom must be informed by state-of-the-art knowledge. The problem has been that such knowledge is dispersed across dozens of disparate disciplines, in thousands of books and journals, with no collation or synthesis to guide best practice. It is no wonder that developers sometimes have to rely on best hunches, rather than best practices, in their design efforts.

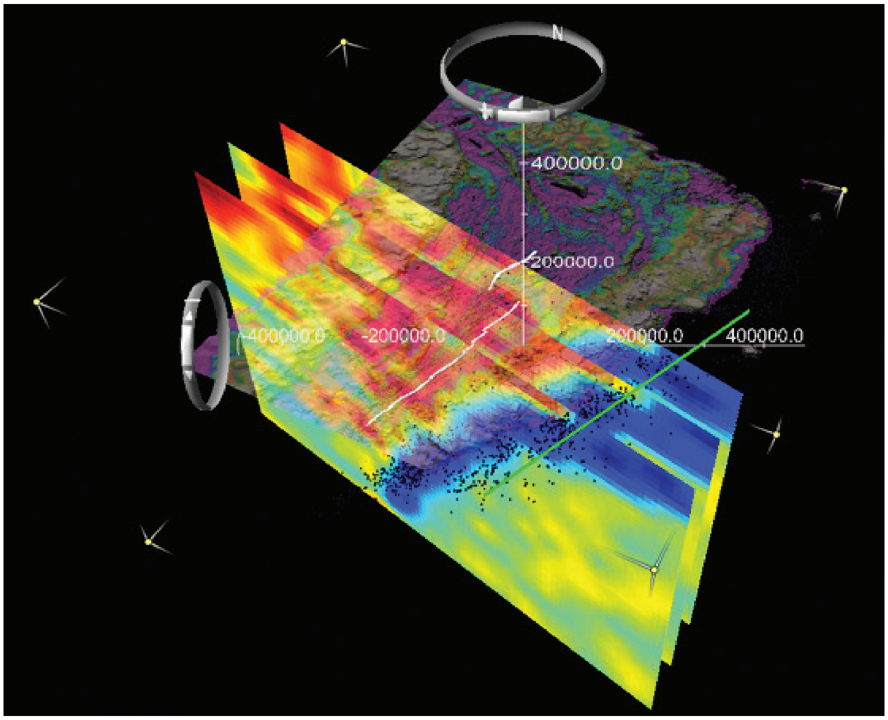


Figure 1. Experts use sophisticated data visualization techniques that may be very difficult for novices to understand. The displayed image is a snapshot from an interactive 3D visualization of the Lau Basin and Tonga Trench curtesy of Allison Jacobs. (Source: SIO Visualization Center, Scripps Institute for Oceanography Institute for Geophysics and Planetary Physics. Retrieved from siovizcenter.ucsd.edu/library/objects/detail.php?ID=138.)

To support interface and tool designers in their efforts to bridge cyberinfrastructure to the classroom, NSF funded Education Development Center, Inc. (EDC), and Scripps Institution of Oceanography to conduct the Oceans of Data project. Our goal has been to identify pertinent literature and expert opinion from the wide-ranging disciplines, to organize that knowledge into an initial integrated framework, to develop considerations and guidelines for educational interface design, and to present them in Visualizing Oceans of Data: Educational Interface Design, a knowledge status report (KSR).

We developed this KSR as a handbook with two key components:

- Guidelines for interface and data visualization tool development
- The considerations (principles, research, and theory) that inform these guidelines

Who Is the Audience for Visualizing Oceans of Data?

Our primary audience for the KSR is developers of interfaces for novice users. These developers will design and create interfaces that are easily navigable. They will define the capabilities that should be built into tools for visual representations of data, be they maps, graphs, or animations. They will construct important functionalities, such as varied color palettes suited to particular purposes, layering of information, alternative formats for representing particular data, and modes for scaffolding to support learning.

A caveat is in order: While the project goal was to array options for interface developers to consider, we recognize that, optimally, design decisions should be made in context—that is, taking into consideration the particular curriculum, the precise learning and teaching goals, and the needs and abilities of particular groups of students. Making appropriate design decisions therefore involves a cast of characters beyond interface developers (see Figure 2). This includes curriculum writers who understand how to guide students in their use of data to meet learning goals, and teachers who play perhaps the most critical role in facilitating students' use of data in the classroom.

Realizing the potential of large databases for student learning also requires the participation of an even wider set of actors. The scientists and database architects who develop the science cyberinfrastructure databases are pivotal. Professional development experts are necessary to help pre-college teachers gain confidence using scientific data and to help them develop strategies for engaging students with this new type of learning activity. Researchers are likewise central in continuing to fill knowledge gaps and build new understandings about learning in this new context. We hope that the KSR will be of interest and assistance to all of these key players as well.

This collaborative project considered in particular the complex observational data that are collected to support scientific research about the earth's oceans, atmosphere, and geosphere. However, the Key Underpinnings and guidelines in this document also have broader application to other scientific domains that hope to support students' access to and visualization of professional scientific databases.

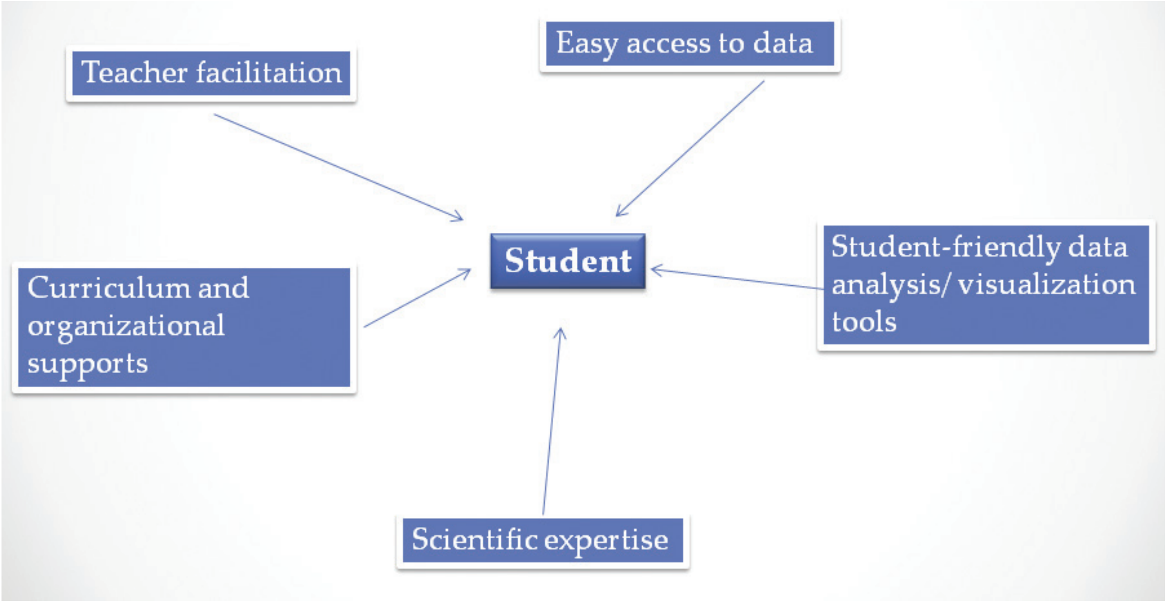


Figure 2. Careful design and testing of each of these elements is required to engage students in scientific practices using data in an online interface

The KSR at a Glance

By summarizing and organizing literature and expert opinion on the tenets underlying design recommendations, as well as the pros, cons, unknowns, and contradictions that sometimes emerge, we created this KSR to inform the process of developing interfaces and tools for data visualizations in the form of georeferenced data representations, graphs, and animations. The KSR is organized as follows:

II. KEY UNDERPINNINGS

Research and theory regarding three topics that are foundational to learning:

- Cognitive Load Theory: The mechanisms of working memory and long-term memory and how they relate to knowledge acquisition
- Visual Perception and Processing: How humans take in and make sense of visual information
- Schemata: How knowledge is stored, organized, and applied

III. CROSS-CUTTING GUIDELINES

Principles and corresponding recommendations that apply across the board to the design of interfaces and data visualizations:

- Adjust Cognitive Load: Designing the presentation of material so that it doesn't exceed the amount of information the learner can actively process
- Draw Attention to Important Features and Patterns: Promoting learning by using methods to highlight key information
- Enable Customization: Building in the capacity to meet different learner needs

IV. SPECIFIC CONSIDERATIONS AND GUIDELINES

The functions and tools particularly relevant to providing access to large scientific databases and facilitating students' work with these data. Design features to be used—or avoided—are addressed for the following:

- Accessing Data: Facilitating the selection and viewing of data parameters
- Geo-Referenced Data Representations (Plan Views, Cross-Sectional Views, and 3D Views): Promoting comprehension and analysis of geographically referenced data visualizations
- Graphs: Supporting interpretation of relationships among data using graphs
- Animations: Using dynamic presentations to represent change over time

V. FUTURE RESEARCH AND DEVELOPMENT: MAPPING THE TERRAIN

Questions relating to the following are presented to map the terrain of research and development that is needed and to focus on certain areas that we believe will be particularly fruitful:

- Authentic Data and Student Learning
- Interfaces and Data Visualization Tools
- Curriculum and Teacher Supports

How to Use the KSR

The KSR serves as both a reference and a tool. It is by no means a step-by-step blueprint for constructing interfaces and tools, for as yet there is no definitive state-of-the-art process for making large scientific databases usable by novice learners. What we offer, rather, is a resource to consult during the software planning and development processes. We know that the considerations and guidelines herein are many and complex. You may choose to pick the low-hanging fruit or to tackle a wide range of approaches. Whatever your *modus operandi,* we do have one recommendation for using the KSR: Please pay heed first to the Key Underpinnings and Cross-Cutting Guidelines chapters, for they offer an abridged orientation to the research, principles, and theories that too often remain under the radar. They also provide a basis for contemplating the considerations and guidelines in the subsequent chapter regarding data access, georeferenced data representations, graphs, and animations.

Students: The Ultimate Beneficiaries

Design decisions must of course be rooted in an understanding of the ultimate user group—students with limited prior experience working with professionally-collected scientific data. Throughout the KSR, we consistently discuss the characteristics and needs of the learners to be served.

The students for whom interfaces and visualization tools will be designed constitute a homogeneous yet diverse user group. Most will be in science classes that stress inquiry and will be called on to engage in key scientific practices, including, for example:

- Asking questions
- Developing and using models
- Planning and carrying out investigations
- Analyzing and interpreting data
- Using mathematics and computational thinking
- Constructing explanations
- Engaging in argument from evidence
- Obtaining, evaluating, and communicating information (National Research Council, 2012)

Virtually all K-16 students will begin their science studies as novices—that is, they will not have the expertise of scientists. As novices they will lack the kinds of knowledge and skill that shape what scientists "attend to and notice, how they organize new information and how they solve problems" (National Research Council, 2006, p. 95). Novices' reasoning and problem-solving will not be fluent. As a whole, they will probably have difficulty drawing inferences from data and making transitions from concrete to abstract thinking. And, of course, all novices will most likely lack any experience whatsoever in working with large science databases.

At the same time, these student users will differ markedly from one another. They will, for example, be divergent in the ways that they most effectively perceive and comprehend information that is presented in a data interface. While some will have more highly developed organizational abilities, some will be less well honed. They will bring different prior knowledge to the class, in terms of science content, mathematical and statistical reasoning, and experience with data visualizations. Their interests and motivation will likewise vary.

Suffice it to say that there is no perfect way to serve all students. But appropriately designed interfaces—in concert with the digital medium's capacity to provide for customization—can go far in igniting students' interest in working with large databases and in supporting their learning.

How We Developed the KSR

How did the notion of the Oceans of Data project arise? How did we go about constructing this resource? Here we describe in broad strokes the path taken . . .

THE INCEPTION

Our collective experience—as science teachers, curriculum developers, designers of student interfaces and curricula keyed to scientific databases, and scientists charged with making a new cyberinfrastructure database accessible to the public—made one thing quite clear: Developers of interfaces that enable nonscientists to work with large databases could use some help in the design process.

The idea of developing a resource to aid developers was exciting, ambitious, and a bit daunting. We marveled at the potential of putting scientists' databases and related tools (in modified forms) into the hands and minds of novice students. We knew that there are few studies of novice use of scientific databases, yet we were familiar with certain bodies of theory and research, as well as observations (our own and others'),that seemed quite germane. And we knew that potentially relevant knowledge was spread across a vast array of fields. Developing the KSR would not be straightforward.

THE PROCESS

From the beginning, we knew that we could not perform the typical literature review/synthesis, where only methodologically rigorous research studies are addressed, because there was so little research regarding access to and use of large scientific databases. We decided on an alternative, though pragmatic, route—addressing theory, expert opinion, and our own experiences, in addition to whatever research existed.

To establish the parameters for our search, we first identified key bodies of knowledge, reviewed some literature, tracked and reviewed some prominent citations in that literature, and conferred with the Oceans of Data Advisory Board and other experts. Thus emerged the focus on two key parameters: the different types of *data representations* that students might encounter (such as georeferenced representations, graphs, and animations), and the processes of *working with data* in which students would likely engage (for example, pattern recognition, finding or selecting data, and reading data representations). Through applying the preliminary coding protocol to several seminal works, we identified a third parameter, dubbed *cross-cutting* issues. This parameter refers to cognitive processes and other factors that relate across the board to various types of representations and actions involved in working with data. The cross-cutting parameters comprise such elements as cognitive load, spatial perception and visualization, prior knowledge, scaffolds and supports, navigation, and schemata. We then established the final coding protocol, while continuing to search for new literature related to our parameters. Testing for inter-rater reliability, we found that the protocol was appropriate to the task at hand and that coders were in agreement.

Our hunt for literature was wide-reaching. We searched a panoply of disciplines, including geosciences education, mathematics education, cognitive psychology, informatics, visual perception, cartography, neuroscience, computer science, learning science, and Universal Design for Learning. We followed up on citations from seminal works in order to ensure that our search was comprehensive and represented the current state of thinking across these fields. All in all, we reviewed over 300 documents (journal articles, books, and presentations), conferred with our ten project advisors, and consulted other experts from a variety of disciplines. We entered articles and other source information into NVIVO software, flagged relevant passages with codes so that we were later able to run queries on individual topics (e.g., animations) and cross-referenced topics (e.g., animations and Cognitive Load Theory) and obtain compilations of relevant quotes. We then summarized the considerations and guidelines that emerged from each query.

Given this burgeoning mass of information from disparate sources, how did we decide what literature to include, guidelines and considerations to report on and how to organize the findings? Following qualitative methods, we noted patterns and themes, identified "disconfirming evidence" (contradictory results), and clustered findings. As we made our judgments, we drew heavily on the collective expertise of the project team:

- 59 years of curriculum development work, including primary authorships of full-year high school Earth science, physics, and chemistry courses
- 20 years in applied science, focused largely on creating and looking for patterns in visualizations of georeferenced data
- 7 years of experience in cognitive science research
- 24 years of research on student learning and pedagogy in science
- 13 years of science teaching in public school classrooms
- 26 years of work in the development of educational software supports for science curricula and computer interfaces to authentic scientific data

Our combined efforts constitute a first step in harnessing knowledge to inform interface development. It is our hope that this KSR will serve as a catalyst for much-needed research, development, and testing so that the field gains a clearer understanding of what design features work (or don't), why, in what contexts, and for whom.

REFERENCES

National Research Council. (2006). *Learning to Think Spatially: GIS as a Support System in the K-12 Curriculum.* Washington, DC: The National Academies Press.

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* Washington, DC: The National Academies Press.

Key Underpinnings

II. KEY UNDERPINNINGS

II. KEY UNDERPINNINGS

This section briefly summarizes a large body of research that is fundamental to understanding how people take in information (such as a map or graph) and make sense of it. These discussions of Cognitive Load Theory, visual perception, and schemata form in large part the basis for the cross-cutting and specific guidelines in the sections that follow.

Cognitive Load Theory

The human brain offers two cognitive structures for storing information:

- *Long-term memory* provides subconscious and permanent storage for practically unlimited amounts of information (Atkinson & Shiffrin, 1968).
- *Working memory* is where information from the environment and/or long-term memory becomes the focus of active attention and processing. Unlike long-term memory, working memory can only hold a finite number of items simultaneously and for a quite limited period of time (Miller, 1956; Peterson & Peterson, 1959). During the learning process, new information is integrated with existing knowledge using working memory resources, and so the way these resources are allocated defines the limits of learning (Paas, Renkl, & Sweller, 2004; Paas, Tuovinen, Tabbers, & Van Gerven, 2003).

Cognitive Load Theory describes three types of demands on working memory:

- *Intrinsic cognitive load* refers to mental effort due to the inherent difficulty of the content to be learned. As the complexity of the content (i.e., the number of interacting elements to be processed) increases, so does the intrinsic cognitive load (Sweller, 1994; Sweller & Chandler, 1994). Visualizations impose increased intrinsic cognitive load when the phenomena or data they represent are complex enough to be challenging to the user.
- *Extraneous cognitive load* refers to any effort required to understand material that's not directly related to the learning process (e.g., mental energy spent trying to find a poorly placed legend on a map). It is of particular concern to interface designers, as extraneous cognitive load often stems from a representation's design or format (Sweller, 1994; Sweller, Chandler, Tierney, & Cooper, 1990). Visualizations imposing extraneous cognitive load require the use of working memory for processing that is not pertinent to the task at hand, thereby reducing the cognitive resources available to engage with new, important, and challenging information.
- *Germane cognitive load* refers to any effort devoted to the construction of new knowledge (Sweller, van Merrienboer, & Pass, 1998). Visualizations that impose germane cognitive load support meaningful engagement with the content and the processing of new information in ways that lead to new or enhanced understandings.

Cognitive Load Theory is an important consideration for those providing students with access to large scientific data sets, such as oceanographic data, and it forms the basis for many of the guidelines in this KSR. Oceanographic and other Earth science data impose a high level of intrinsic cognitive load due to the number of interacting elements typically involved in Earth systems. As a result, it is critical that interface designers take steps to reduce extraneous load, alleviate intrinsic cognitive load, and maximize germane cognitive load.

A key point is that expert scientists already have well-formed domain knowledge in their long-term memory that they can apply automatically, freeing up the necessary working memory resources to read and interpret complex data representations (Kalyuga, Chandler, & Sweller, 1998; Pass & van Merrienboer, 1994; Sweller, 1994). However, novice learners must devote much more of their working memory to knowing how to approach the task, making sense of unfamiliar data sets and visualization formats, and constructing new understandings from what they see. Interface developers need to provide visualizations and other interface features designed to minimize processing demands that are unrelated to content by eliminating unnecessary distractions and providing scaffolding around challenging representation elements. It is also important to note that anxieties felt by the user, whether they stem from an overwhelming and unfamiliar interface or from the user's beliefs about their own competence, can consume additional working memory resources .

So how does one maximize the amount of working memory that is available to find relevant data on a website and/or to identify significant patterns in data visualizations? The discussions of visual perception and schemata that follow shed light on how our eyes and brains work to separate the "signal" from the "noise," and how this information is useful to the design of interfaces to large scientific data sets.

REFERENCES

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2). New York, NY: Academic Press.

Kalyuga, S., Chandler, P., & Sweller, J. (1998, March). Levels of expertise and instructional design. *Human Factors, 40*(1), 1-17.

Miller, G. A. (1956). Human memory and the storage of information. *IRE Transactions on Information Theory*, *2*(3), 129-137.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. Educational Psychologist, 38, 1-4.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38,* 63-71.

Paas, F. G. W. C., & van Merrienboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*(4), 351-371.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193-198.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295-312.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*, 18-233.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *Journal of Experimental Psychology: General, 119,* 176-192.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review,* 10, 251-296.

Visual Perception and Processing

WHAT IS VISUAL PERCEPTION, AND WHY IS IT IMPORTANT?

> *It has become clear that human vision is highly variable and difficult to predict, particularly when the visual task involves a stimulus as complex as a map.*
>
> *—Alan M. MacEachren, How Maps Work, 2004 (p. 22)*

The human visual perception system is not like a camera—our eyes do not take snapshots that simply become "pictures in the head." Visual perception is information-processing, shaped and mediated not only by the sophisticated neural mechanisms of our eyes and brain, but also by cognitive resources in our short- and long-term memory (Mathewson, 1999; Ware, 2000).

The components of the human visual system and associated cognitive processes are highly specialized and have evolved in response to survival demands of the three-dimensional world humans have lived in for thousands of years. Because the use of two-dimensional representations, such as maps and graphs, and the use and navigation of Web interfaces has developed quite recently in human history, our visual perception system is not specifically adapted to these tasks. Therefore, it's critical to understand how to design two-dimensional media to take advantage of the strengths of our highly evolved and complex visual system and to compensate for its weaknesses. Looking at the design of data interfaces through this lens helps us understand, for example, why red stands out (finding the ripe berries in a bush), why movement grabs our attention (hunting and avoiding predators), and why variations in light luminance and shading work better than variations in color hue for perceiving shape and form.

HOW DOES THE HUMAN VISUAL PERCEPTION SYSTEM WORK?

Sensory inputs based on relative differences in light levels prompt cells in the retina (rods and cones) to send information about visual properties, such as color and shape, for further processing in an area of the brain called the visual cortex. Here, neurons attuned to specific areas of the visual field and to certain configurations of visual properties (e.g., contrast, vertical lines, motion) begin forming a rudimentary "sketch" of perceived spatial properties ("where") and form ("what") (Kosslyn & Koenig, 1992; Marr, 1982; Mishkin & Ungerleider, 1982; Plass, Homer,

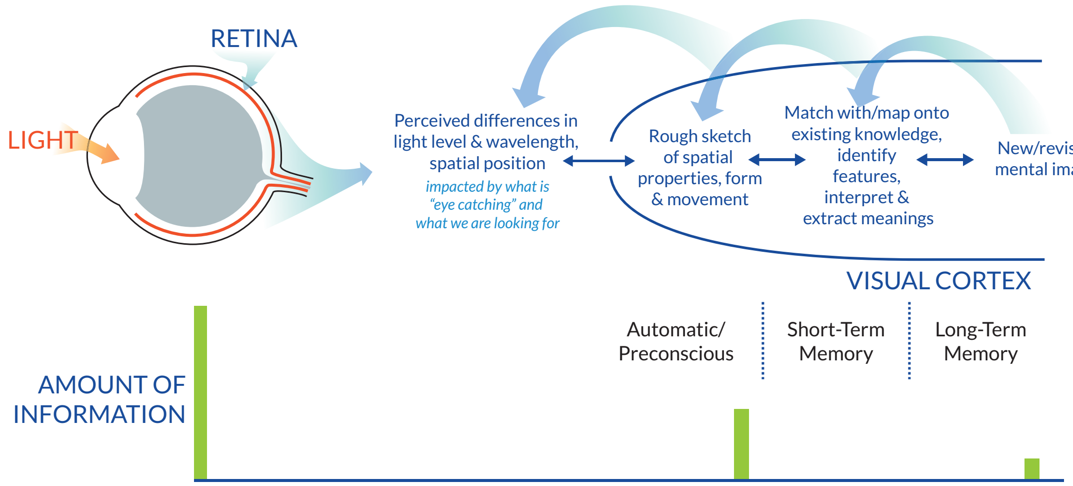


Figure 3. A simplified diagram of the information-processing performed by the human visual system.

& Hayward, 2009). These rough sketches undergo further processing in specialized areas of the visual cortex, where they are mapped on to knowledge we hold in our long-term memory as we identify objects, interpret and extract meanings, and ultimately incorporate these into new or existing understandings (MacEachren, 2004). Figure 3 provides a simplified illustration of this process.

A key point about our visual system relevant to interface designers is that we are surrounded by overwhelming amounts of sensory information with varying degrees of usefulness and interest, depending on the context. Because we don't have the cognitive resources (or the need, really) to capture and consider everything in fine detail, filtration of this information occurs at all levels of the perception process—from stimulation of retinal cells to integration into long-term memory. When we view a complex visual display, we selectively take in and process the information so that the resulting mental image will contain a much smaller, selected set of information (Desimone & Duncan, 2005; Plass et al., 2009, Ware, 2008). The form, sophistication, and accuracy of the mental image that results is mediated to a significant extent by what we already know (i.e., have stored in our long-term memory) (Arnheim, 1969; Barry, 1997; Plass et al., 2009). Therefore, what a novice "sees" in a particular graphic representation can be very different from what an expert sees.

The structure and function of our visual system also determines what we are able to see and how we find and focus on features and patterns. Receptors that trigger our perception of color (cones) are located in the center of the retina, the part that is directly exposed to light entering the eye. Our visual acuity decreases significantly from the center of our vision to the periphery, which means that the colors and details that are visible when we look directly at something will not be clear in our peripheral vision (Johnson, 2010; MacEachren, 2004). When we search for something on which to focus our attention, we use our peripheral vision, so important features in a visualization or on a webpage must be designed to stand out in the periphery.

AUTOMATIC PROCESSING

To ensure efficient processing of the most relevant information at any given moment, we make both preconscious (automatic) and conscious determinations about where to direct our attention and what information to extract. A visual processing architecture designed to support successful navigation and functioning in our environment, along with our cumulative knowledge, experiences, and expectations, determines what we "see"—what stands out, and what fades into the background and essentially becomes "noise." Visual features that are intrinsically eye-catching or that can be recognized and matched with existing schemata without conscious direction are considered "automatically" processed (Plass et al., 2009; Ware, 2008). Because automatic processing does not tax limited working memory resources, the design of a data interface should maximize the amount of information that can be processed automatically. Novices will not have the extensive background knowledge that allows experts to automatically seek out and recognize certain features and patterns. Students' visual perception systems will rely more heavily on the natural tendencies of the human visual system to automatically recognize and prioritize certain shapes, colors, and forms; to interpret visual information in certain ways; and to group certain elements of a static or animated display (Ware, 2000). This means that novices are likely to notice different features than experts do and to therefore build a very different mental image. Awareness of these natural biases is critical to developing an interface design that focuses students' attention appropriately, maximizes their engagement with germane cognitive load, and minimizes the degree to which they may draw wrong conclusions and incorporate misinformation into their mental image (Lowe, 2003; Shah, 2002).

Research on visual perception and processing suggests that the following features, used often in data representation, are processed automatically:

- Form features (e.g., line length, line orientation, size, spatial grouping)
- Color (hue and intensity)
- Motion (flickering or motion on a trajectory)
- Spatial position (e.g., 2D position or stereoscopic depth) (Ware, 2000)

It is important to note that in general, the human visual system is sensitive to abrupt changes, contrast, or disruptions in visual features, such as those listed above. For example, we are more likely to attend to and perceive a sharp divide between a black region and a white region than the gradual fade from black to gray to white (T. Shipley, personal communication, November 1, 2012).

GROUPING

During the initial stage of visual processing, our visual system automatically groups elements into objects, shapes, and categories that can be mapped onto those that are in our long-term memory (MacEachren, 2004). The automatic grouping and matching that occurs differs from one person to another, depending on the person's level of prior knowledge and experience with what is being viewed, where those with more experience may group features more accurately and appropriately. However, there are certain basic tendencies built into human visual perception and cognition that are important for interface designers to consider. For example, the physical arrangement of elements in a visualization can convey certain information (such as what category features belong to, or how much of something there is) without requiring focus and active thought from the viewer, and thus preserve valuable short-term memory resources (Ware, 2000). On the flip side, if the groupings are inappropriate, the viewer may jump to the wrong conclusions because of these natural tendencies, and as a consequence may misinterpret or struggle to understand the information presented.

The Gestalt school of psychology conducted some of the first research on pattern perception and produced a set of design principles that describe how humans perceive patterns and relationships between objects and groups of objects. A number of these principles are relevant to the design of data interfaces and visualizations:

- The principle of *proximity* or *contiguity* states that things that are closer together will be seen as belonging together.
- The principle of *similarity* states that things that share visual characteristics, such as shape, size, color, texture, value, or orientation, will be seen as belonging together.
- The principle of *continuity* predicts the preference for continuous figures.
- Related to the principle of *good continuation*, we have a tendency to close simple figures, independent of continuity or similarity. This results in an effect of filling in missing information or organizing the information that is present to make a whole.
- The principle of *area* states that the smaller of two overlapping figures is perceived as figure while the larger is regarded as *ground.*
- The principle of *symmetry* describes the instance where the whole of a figure is perceived rather than the individual parts that make up the figure.

There are also rules that are relevant to dynamic displays. For example, objects moving together are seen as a group. Also, if we initially perceive objects as being part of a group, our perception will try to maintain a stable state—even if their position changes, we will try to retain the initial group. It's particularly important to note that familiar shapes or arrangements form groups (Koffka, 1935).

Over time and with practice, novice students will learn to recognize certain patterns in scientific data visualizations and will begin to process them automatically, as experts do.

COLOR

Color is a valuable tool that can be used to direct visual attention to significant features and patterns. Specialized receptors in the retina are sensitive to stimuli that trigger our perception of colors (i.e., light of different wavelengths), and our perception of color helps us make unique judgments about our environment. Color is processed and interpreted somewhat differently from other visual features, such as size or shape, and for the most part is not critical to successfully navigating day-to-day activities. However, color does serve important evolutionary functions, such as breaking camouflage and providing clues about an object's physical properties (Ware, 2000).

Our visual processing system detects variations along three dimensions that are often associated with color and are of importance to effective visualization design: hue, luminance, and saturation. When these dimensions are applied in ways that respond to how our visual system works, they can be powerful tools for representing variations in data sets, assisting in the grouping of certain elements or features, or calling attention to important features of a data visualization or website:

- *Hue* refers to what we typically think of as colors, such as purple, yellow, or blue, which are our visual system's interpretation of the patterns of light wavelengths that are reflected toward our eyes.
- *Luminance* refers to the amount of visible light. While luminance is a measurable attribute, the terms "lightness" and "value" are also used to describe the perceived brightness of a color. It's important to note that our eyes perceive contrast more than the total amount of light, which makes sense since we experience different light levels over the course of a day (MacEachren, 2004).
- *Saturation* is a scientific term used to describe how vivid a color appears to the viewer. High saturation colors are seen as more intense or pure, while low saturation colors are seen as being closer to black, white, or gray.

Variations along these dimensions influence what catches our eye, our perception of three-dimensional shapes and depth in a scene, and our perception of magnitude (e.g., more or less, higher or lower) (Ware, 2000). However, these perceptions may or may not be accurate, depending on how color is used in a visualization. For example, hue is useful for identifying or categorizing objects, but luminance and saturation better convey magnitude, three-dimensional shape, and depth (Rogowitz & Treinish, 1996).

Our visual system is susceptible to certain illusions when viewing a color image as well. For example, the perceived size of an object or area may be influenced by its color, with red or purple objects or regions on a map appearing larger than blue or green objects (Cleveland & McGill, 1983; Rheingens & Landreth, 1995; Tedford, Berquist, & Flynn, 1977); perceived hue can be influenced by saturation and vice versa; the perceived depth of an object or area can be influenced by its hue (e.g., red can look closer than green when the two are put next to each other); and perceived hue, luminance, or saturation can be influenced by the color of surrounding objects (Harrower & Brewer, 2011; Rheingens & Landreth, 1995).

ATTENTION-GRABBERS

Certain features stand out and grab our attention, and are more likely to be perceived automatically and thus elevated in importance relative to other features. For example, movement grabs our attention, as do certain colors, such as red and yellow (Ware, 2000). These "pop out" features are very important, and therefore are discussed and incorporated into our Cross-Cutting Guidelines on drawing attention to important features and patterns.

It's important to recognize that while some features automatically "pop" and help organize visual input into meaningful shapes, objects, and spatial relationships, visual perception is not a one-way street. Pre-existing knowledge, experiences, and ideas influence not only our interpretations of visual information, but also where we direct our attention, what features we extract, and which objects and patterns we recognize automatically, with minimal use of working memory (MacEachren, 2004). One goal for developers should be to support the construction of robust schemata (i.e., a framework of relevant knowledge and skills), so that users can engage with increasingly complex data visualizations and tools. The next section, thus, focuses on schemata—what they are and how they operate.

REFERENCES

Arnheim, R. (1969). Art and visual perception: *A psychology of the creative eye.* Berkeley, CA: University of California Press.

Barry, A. M. (1997). *Visual intelligence: Perception, image, and manipulation in visual communication*. Albany, NY: SUNY Press.

Cleveland, W. S., & McGill, R. (1983). A color-caused optical illusion on a statistical graph. *American Statistician, 37*, 101-105.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193-222.

Harrower, M., & Brewer, C. A. (2011). ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader: Theories of Mapping Practice and Cartographic Representation.* Chichester, UK: John Wiley & Sons, Ltd. doi: 10.1002/9780470979587.ch34

Johnson, J. (2010). *Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules.* Burlington, MA: Morgan Kaufmann.

Koffka, K. (1935). *Principles of Gestalt Psychology.* London, England: Routledge & Kegan Paul.

Kosslyn, S. M., & Koenig. O. (1992). Wet mind: *The new cognitive neuroscience.* New York, NY: Free Press.

Lowe, R. K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction, 13*(2), 157-176.

MacEachren, A. (2004). *How Maps Work.* New York, NY: Guilford Press.

Marr, D. (1982). *Vision.* San Francisco, CA: W. H. Freeman.

Mathewson, J. H. (1999). Visual-spatial thinking: An aspect of science overlooked by educators. *Science Education, 83,* 33-54.

Mishkin, M., & Ungerleider, L. G. (1982). Object Vision and Spatial Vision: Two Cortical Pathways. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 296-302). Cambridge, MA: MIT Press.

Plass, J. L., Homer, B. D., & Hayward, E. (2009). Design factors for educationally effective animations and simulations. *Journal of Computing in Higher Education, 21*(1), 31-61.

Rheingans, P., & Landreth, C. (1995). Perceptual Principles for Effective Visualizations. In G. Grinstein & H. Levkowits (Eds.), *Perceptual Issues in Visualization* (pp. 59-74). Berlin, Germany: Springer-Verlag.

Rogowitz, B. E., & Treinish, L. A. (1996, May/June). How Not to Lie with Visualization. *Computers In Physics, 10*(3), 268-273.

Shah, P. (2002). Graph comprehension: The role of format, content and individual differences. In M. Anderson, B. Meyer, & P. Olivier (Eds.), *Diagrammatic Representation and Reasoning.* Berlin, Germany: Springer.

Tedford, W. H. Jr., Berquist, S. L., & Flynn, W. E. (1977). The Size-Color Illusion. *The Journal of General Psychology, 97,* 145-149.

Ware, C. (2000). *Information Visualization: Perception for Design.* San Francisco, CA: Morgan Kaufmann.

Ware, C. (2008). *Visual Thinking for Design.* Burlington, MA: Morgan Kaufmann.

Schemata

Having a basic understanding of how knowledge is organized and applied can help shed light on what makes scientific visualizations challenging and, when poorly designed, ineffective for novice users. However, the cognitive processes involved are complicated and not easily or consistently defined, even by those who study them. *In How Maps Work* (2004), MacEachren describes this challenge:

> *These structuring mechanisms have been given a variety of labels depending upon the kind of knowledge they are hypothesized to deal with and the proclivities of the author suggesting them (e.g., schemata, frames, scripts, mental models, idealized cognitive models, etc.). Although there are some fundamental differences implied by these labels (and among authors who use the same label), for simplicity in discussing how map understanding may depend upon these cognitive structures I will adopt a single term, schemata, to refer to them. (p. 174)*

Like MacEachren, we have opted in this KSR to primarily use the term *schemata* (with occasional deference to authors' language choices) to refer to organizing frameworks that provide a format for representing categories of "whole" objects, systems, processes, or domains by characterizing the relationships that connect the various parts or elements (Greeno, 1989; Johnson-Laird, 1983). Serving both explanatory and predictive purposes, *schemata* are the lenses through which we interpret our surroundings and experiences and the frameworks we apply when drawing on existing knowledge to approach a particular task or problem.

From birth, humans begin forming schemata by categorizing and organizing information about the world around them. As we gain more and more experience related to a particular schema, it becomes a more robust, coherent, and elaborate part of our long-term memory. In addition, as we become more experienced, our schemata become more automated and require fewer cognitive resources to apply. A fully automated schema can be activated and applied without conscious direction, allowing for more automatic processing of the features, patterns, and configurations we know well, based on prior knowledge and experience.

The letters of the alphabet provide a simple, concrete, and familiar example:

k *k* **k** k K *K*

Each letter above looks quite different, but it's immediately obvious that they are all the letter "K," because our schema for this letter has been evolving since the beginnings of our literacy to accommodate different fonts, sizes, colors, orientations, etc.

Consider another example that illustrates the range of purposes and grain sizes that schemata can have. Prior to leaving for work in the morning, you might check the local weather forecast. A quick investigation of the weather map in Figure 4 shows large areas of dark green over your town. Drawing on your well-developed schema for weather maps, you quickly conclude that it's raining in your area and begin to plan accordingly. Then, drawing on your schema for a rainy day, you conclude that you should wear rain boots, bring an umbrella to work, and expect a slightly longer commute.

Schemata are maintained in the virtually unlimited store offered by long-term memory and are continuously in flux as they adapt to incorporate new information. When applied, schemata allow for the "chunking" of interacting or related elements into individual units in working memory, freeing cognitive resources for engaging with the task at hand (Chi, Glaser, & Rees, 1982; Sweller & Chandler, 1994).

Interface designers should care about schemata for three primary reasons:

- Novice users lack the robust and automated schemata that experts use to interpret and manipulate data visualizations. Because schemata both (1) tell users what to look for and how to interact with the display, and (2) operate as single entities in working memory, this means that experts have the capacity to engage more quickly and extensively with data visualizations and to identify increasingly complex patterns and relationships (Chi et al., 1982; Sweller & Chandler, 1994).

- When used appropriately, data visualizations can activate users' existing, but perhaps not fully automated, schemata (i.e., prior knowledge), reducing the need for conscious processing and demands on working memory.
- Poorly designed data visualizations can lead to the construction of inaccurate or incomplete schemata (i.e., misconceptions), setting up users for future difficulties down the road.

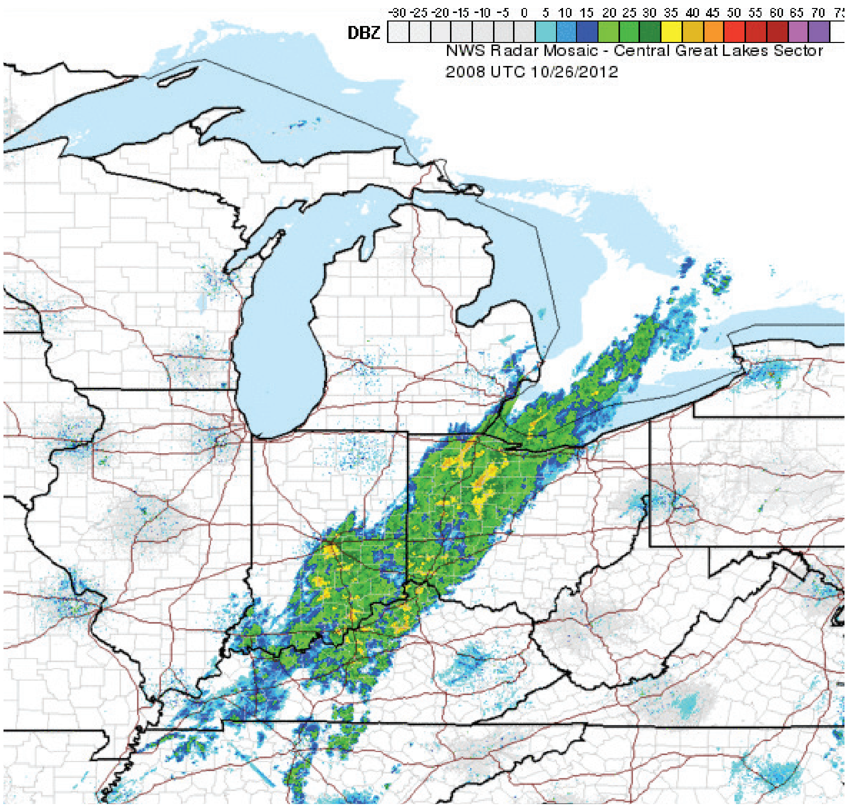


Figure 4. Chances are good that your robust and automated schemata allow you to rapidly interpret this visualization of the weather in your area. (Source: National Weather Service Enhanced Radar Mosaic. Retrieved from radar.weather.gov/ridge/Conus/centgrtlakes.php.)v

Consider how the weather map example above might have played out differently if the individual in question were a nine year old getting ready to play outside rather than an adult on his or her way to work. How helpful would the same map have been? Assuming that the nine year old could make any sense of the map, how much more mental energy would he or she have to use in trying to interpret it? Interface designers should entertain similar questions when considering the gap in expertise between expert scientists and more novice users in science classrooms, and what the implications might be for visualization design. Automated schemata reduce the burden on working memory and facilitate quick search and recognition. Scientists, having ample experience (and thus more automated schemata) both in the content area and in interpreting and manipulating data representations, will require fewer cognitive resources to scan the visualization, interact with the visualization's controls, quickly identify patterns (or the lack thereof), and interpret their meaning. Meanwhile, the novice user's schemata might provide little help on any of these fronts, mostly likely leading to an overwhelmed working memory.

Finally, it is important to highlight the processes that influence how schemata are applied. The information we gather from automatically processed features helps us determine which schemata are most appropriate for the task at hand and therefore which to activate (MacEachren, 2004). Interface designers should consider this when making decisions around visualization format and design.

The guidelines in the sections that follow will help designers avoid fostering inaccurate schemata by making sure that users can identify the appropriate features and patterns without overburdening their working memory.

REFERENCES

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7-75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Greeno, J. G. (1989). Situations, mental models, and generative knowledge. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 285-318). Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

MacEachren, A. M. (2004). *How Maps Work: Representation, Visualization, and Design.* New York, NY: The Guilford Press.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*, 185-233.

Cross-Cutting Guidelines

III. CROSS-CUTTING GUIDELINES

III. CROSS-CUTTING GUIDELINES

The guidelines in this section apply broadly to the design of data interfaces for novices. The next section offers guidelines specific to accessing and selecting data and to working with data displayed as georeferenced data representations, graphs, and animations.

Cross-Cutting Guideline 1: Adjust Cognitive Load

As discussed in II. Key Underpinnings: Cognitive Load Theory, the finite capacity of working memory limits the amount of information that the human brain can actively attend to at one time; for novices, it's particularly important that interface and visualization design support the most effective use of the limited resources available. This can be done by maximizing the germane cognitive load and minimizing both the extraneous cognitive load and, to the extent possible, the intrinsic cognitive load associated with students' work with the data interface, as shown in Figure 5.

While an interface designer may have little to no control over the content's inherent difficulty, there are ample opportunities to make design choices that reduce the number of cognitive resources needed to search, manipulate, and process the information presented in a data interface. While taking steps to reduce extraneous cognitive load frees up resources that can then be devoted to learning, novices can't integrate this new information into their existing knowledge unless germane cognitive load is sufficient. Further research is needed on designs that actively increase germane cognitive load.

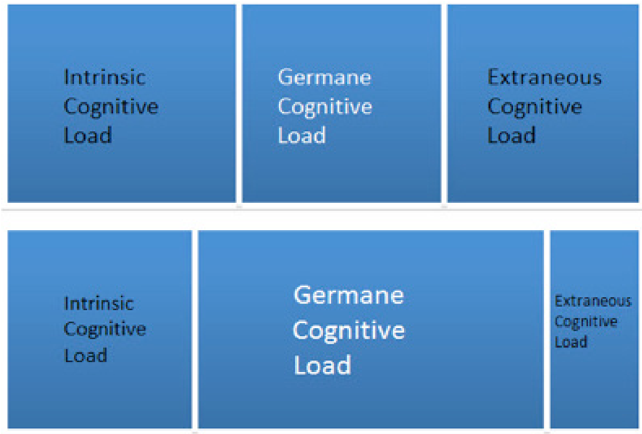


Figure 5. To maximize the learning that takes place when a student uses a data interface, it's important to maximize germane cognitive load and minimize extraneous and intrinsic cognitive load.

PROVIDE COMPLEMENTARY INFORMATION IN MULTIPLE FORMATS

Evidence suggests that different kinds of sensory input (e.g., visual or auditory input) go through unique processing channels as we make meaning from raw, physical signals and that these processing channels draw on partially independent stores of working memory (Mayer & Anderson, 1991; Mayer & Moreno, 1998; Moreno & Mayer, 1999). Paivio's dual-coding theory (1986), Baddeley's model of working memory (1992), and Mayer's cognitive theory of multimedia learning (2005) argue that auditory and visual information are processed in distinct channels, each with their own working memory capacities, and that using both channels could not only enhance learning but also allow more information to be processed simultaneously. This is not to suggest, however, that representing verbal information both visually and as narration is always a good idea. Depending on the user and the content in question, it may only be helpful in instances where neither the graphic portion nor the text portion of the visualization can be fully understood on its own.

INTEGRATE TO FOCUS USER ATTENTION

While providing multiple information formats may alleviate some of the intrinsic cognitive load imposed by complex and challenging content, if not done properly, any gains might be lost as users spend additional working memory resources going back and forth between information sources that are poorly integrated (i.e., located in different parts of the visualization) (Cook, 2006; Hegarty & Just, 1993; Jeung, Chandler, & Sweller, 1997;

Mayer & Moreno, 2003; Moreno & Mayer, 1999; Vekiri, 2002). For multiple modes of input to be effective, they must be coordinated in space (e.g., placing text within a graphic visualization as opposed to underneath it or in a different window) and time (e.g., timing narrated content so that it corresponds to what's currently happening in an animated visualization) (Mayer & Gallini, 1990). This prevents users from having to hold partially complete information in their working memory until they can find, process, and integrate supplementary information presented elsewhere or at another time. Integrating different modes of information presentation is particularly critical when novices are presented with challenging content that imposes high levels of intrinsic cognitive load (Chandler & Sweller, 1991, 1992, 1996; Cook, 2006; Mayer & Anderson 1991, 1992; Mayer & Gallini, 1990; Sweller, Chandler, Tierney, & Cooper, 1990; Tarmizi & Sweller, 1988; Ward & Sweller, 1990).

RECOGNIZE HURTFUL (AND HELPFUL) REDUNDANCIES

Depending on a user's expertise and the difficulty posed by the visualization, redundant information (such as text describing the content of the visualization, or labels on self-evident features) could provide a critical support for novices to engage meaningfully with the material. However, it could also impose unnecessary demands on working memory if the user has enough background knowledge or if the material is straightforward enough to be interpreted without supportive text. Processing the same information presented in a different format requires additional cognitive resources without adding any new information to enhance understanding. In these cases, redundant information should be omitted. When considering including text supports, it's helpful to ask: *Will users be able to adequately interpret meaning from the graphic representation alone? I*f so, avoid the inclusion of redundant information. If not, providing text or other representations of the same content will be beneficial (Cook, 2006; Hasler, Kersten, & Sweller, 2007; Sweller, 2002, 2004).

ELIMINATE UNNECESSARY DISTRACTIONS

In addition to omitting features that are informationally redundant and unhelpful to users, eliminating visualization elements that are purely decorative will further reduce extraneous cognitive load. There is often a tendency to include extra "bells and whistles" to make visualizations "fun." However, evidence suggests that if these extra features are unrelated to the task at hand and do not contribute to increased understanding of the material, they run the risk of posing undue demands on working memory (Moreno & Mayer, 2000).

WORK WITH VISUAL PROCESSING MECHANISMS

The human visual system has evolved to help us survive and navigate in our three-dimensional world. The resulting biases of our visual processing system can be used to the novice's advantage in order to reduce the amount of working memory resources devoted to searching or scanning a visualization for important features and patterns and mapping visualization content onto existing schemata. (See Cross-Cutting Guideline 2: Draw Attention to Important Features and Patterns for an explanation of how to use visual perception system biases to draw attention to important information.)

CONSIDER THE NUMBER OF REPRESENTATIONS PRESENTED SIMULTANEOUSLY

Multiple representations allow students to explore data from different perspectives and to use different strategies, which can help them develop more robust understandings (Seufert, 2003). However, many studies have shown that it is very difficult for users to translate between representations (Ainsworth 2006). Coordinating multiple data representations is cognitively demanding and can overwhelm the working memory capacity of novice users, as they need to study and understand each representation individually, discern how they are related to each other, and finally relate both to the underlying scientific concept (Ainsworth, 2006; Gilbert, Reiner, & Nakhlel, 2008; Seufert, 2003). Some students will expend much of their cognitive resources simply interpreting the individual graphics and won't have the capacity to link them (Cook, 2006; Yerushalmy, 1991). Novices also tend to focus on the representation that is simplest, most concrete, and/or most familiar (Cook, 2006; Cox

& Brna, 1995, Scanlon, 1998; Tabachnek & Simon, 1998), which may prevent them from grappling with the more challenging but potentially more germane representation. Given the challenges faced by learners, research suggests that the decision to present more than one representation at a time should be considered carefully in light of the students' prior knowledge and the requirements of the task they will perform.

PROVIDE FLEXIBLE SUPPORTS TO OPTIMIZE COGNITIVE LOAD FOR USERS WITH A RANGE OF EXPERIENCE AND KNOWLEDGE

One size does not fit all when it comes to designing for optimal levels of cognitive load. The reviewed literature suggests a number of strategies (such as the guidelines offered in this section) for reducing the extraneous and intrinsic cognitive load that a novice experiences when confronted with new and complex material, while at the same time ensuring the development of accurate new understandings. However, using the same strategies with more expert users might actually reduce the visualization's effectiveness. More knowledgeable and experienced users are able to extract patterns and meaning with fewer supports, such as explanatory text. For these users, adding extra supports increases the amount of unnecessary processing and decreases the working memory resources available for productively processing a visualization's content (Kalyuga, Ayres, Chandler, & Sweller, 2003; Kalyuga, Chandler, & Sweller, 1998).

REFERENCES

Ainsworth, S. E. (2006). DeFT: A conceptual framework for learning with multiple representations. *Learning and Instruction, 16*(3), 183-198.

Baddeley, A. (1992). Working Memory: The Interface between Memory and Cognition. *Journal of Cognitive Neuroscience, 4,* 281-288.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293-332.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology, 62,* 233-246.

Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Psychology, 10*, 151-170.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education, 90*(6), 1073-1091.

Cox, R., & Brna, P. (1995). Supporting the use of external representations in problem solving: The need for flexible learning environments. *Journal of Artificial Intelligence in Education,* 6, 239-302.

Gilbert, J. K., Reiner, M., & Nakhlel, M. (Eds.). (2008). *Visualization: Theory and Practice in Science Education.* New York, NY: Springer.

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. Applied Cognitive Psychol*ogy, 21*, 713-729.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32,* 717-742.

Jeung, H., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology, 17,* 329-343.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*(1), 23-31.

Kalyuga, S., Chandler, P., & Sweller J. (1998). Levels of expertise and instructional design. Human Factors, *40,* 1-17.

Mayer, R. E. (Ed.). (2005). *Cambridge handbook of multimedia learning.* New York, NY: Cambridge.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83,* 484-490.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology, 84,* 444-452.

Mayer, R., & Gallini, J. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 88,* 64-73.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90,* 312-320.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. In R. Bruning, C. A. Horn, & L. M. PytlikZillig (Eds.), *Web-based learning: What do we know? Where do we go?* (pp. 23-44). Greenwich, CT: Information Age Publishing.

Moreno, R., & Mayer, R. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91,* 358-368.

Moreno, R., & Mayer, R. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology, 92*(1), 117-125.

Paivio, A. (1986). *Mental representation: A dual coding approach.* New York, NY: Oxford.

Scanlon, E. (1998). How beginning students use graphs of motion. In M. van Someren, P. Reimann, H. P. A. Boshuizen, & T. de Jong (Eds.), *Learning with multiple representations* (pp. 67-86). Oxford, England: Elsevier.

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction, 13*, 227-237.

Sweller, J. (2002). Visualisation and instructional design. In R. Ploetzner (Ed.), *International workshop on dynamic visualizations and learning.* Tubingen, Germany: KnowledgeMedia Research Center.

Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science, 32*, 9-31.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structure of technical material. *Journal of Experimental Psychology: General, 119,* 176-192.

Tabachnek, H. J. M., & Simon, H. A. (1998). One person, multiple representations: An analysis of a simple, realistic multiple representation learning task. In M. van Someren, P. Reimann. H. P. A. Boshuizen, & T. de Jong (Eds.), *Learning with multiple representations* (pp. 197-236). Oxford, England: Elsevier.

Tarmizi, R., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80,* 424-436.

Vekiri, I. (2002). What Is the Value of Graphical Displays in Learning? *Educational Psychology Review, 14(*3), 261-312. Retrieved from http://hdl.handle.net/2027.42/44453

Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction, 7*, 1-39.

Yerushalmy, M. (1991). Student perceptions of aspects of algebraic function using multiple representation software. *Journal of Computer Assisted Learning, 7*, 42-57.

Cross-Cutting Guideline 2: Draw Attention to Important Features & Patterns

Visualizations offer opportunities to investigate complex phenomena and systems using large and "messy" data sets. The visualization format, the nature of the data sets under investigation, and key content-related understandings will all be quite familiar to experts, who have fully developed and robust schemata related to the phenomena they're investigating, the patterns they're looking for, and the visualization formats and controls used to surface those patterns (Cook, 2006; Germann & Aram, 1996; Lowe, 2003) and will likely have little trouble identifying important patterns or features. However, as described in II. KEY UNDERPINNINGS, novices who lack these schemata will perceive complex visualizations in a very different way.

When processing any visual information from our environment, the details that ultimately form our mental image are determined by their location in the visual field, our pre-conscious processing of features automatically detected by the human visual system, and our conscious decision to attend to certain areas or details of the display based on our prior knowledge and goals (Mathewson, 1999; Ware, 2000).

PLACE IMPORTANT FEATURES IN THE CENTER OF THE FIELD OF VIEW

The human eye is sensitive to light from a wide field of view; however, it is important to recognize that the amount and kinds of information processed will vary depending on where in the visual field a stimulus falls. Highest levels of detail are detected at the center of the visual field, and the ability to detect fine details quickly drops off as you move farther toward the peripheries. It's helpful to think in terms of a "useful" field of view—the area in which visual features can be rapidly processed. The size of the useful field of view varies based on the density and granularity of the information presented (Ware, 2000). For example, a user's attention will focus more narrowly on a scatter plot that is densely populated with data points than one whose data points are sparser and more spread out. Research has illustrated the impact of "field of view" on participants' ability to search and identify features in a visualization; participants were able to locate targets at the center of displays (including maps) more quickly than those in peripheral locations. In addition, targets were more easily found when they were located toward the top of a display versus the bottom, and on the right side of a display versus the left (Brennan & Lloyd, 1993; Lloyd, 1997; Wolfe, Klempen, & Dahlen, 2000). When designing an interactive display, it's important to be aware of where the user's attention was most recently focused so that important next steps can be placed in close proximity (Johnson, 2010).

The ability to accurately perceive colors also drops off in the peripheries of the human field of view, as cones—the retinal cells sensitive to light wavelengths that our brain interprets as colors—are concentrated in the center of the eye. Interface designers should ensure that critical features and patterns, especially those coded using color or featuring fine detail, are located near the center of a display (Lloyd & Bunch, 2003; Ware, 2000).

MAKE IMPORTANT FEATURES & PATTERNS DISTINCT FROM THE BACKGROUND

To appropriately focus novice users, it is important to avoid inadvertently highlighting background or unimportant elements of a display and to strategically cue conscious attention to features and patterns that might otherwise be missed or misinterpreted by novices (Rheingens & Landreth, 1995). Doing so will alleviate demands on limited working memory resources by reducing unsystematic searching and scanning, and will help to ensure the development of accurately enhanced mental models (Lloyd, Hodgson, & Stokes, 2002; Rieber, 1991). Evidence suggests that objects that are brightly colored, moving or changing, defined by sharp boundaries, or highly saturated will likely catch the eye (Rheingans & Landreth, 1995). However, research also shows that it's often the relative differences between a target feature and potentially distracting elements that most affect the search process (Johnson, 2010; Ware, 2000, 2008).

APPLY GESTALT PRINCIPLES TO CLEARLY REPRESENT PATTERNS & RELATIONSHIPS

In II. KEY UNDERPINNINGS: VISUAL PERCEPTION AND PROCESSING, we introduced principles of Gestalt theory that describe many of the ways that humans tend to group objects based on visual input. Interface designers should make visual relationships clear in accordance with these principles. For example, related items in a visualization should be visually "chunked" by placing them near each other and/or by assigning similar visual attributes, such as shape, size, color, texture, value, or orientation. Johnson (2010) says, "A recommended practice, after designing a display, is to view it with each of the Gestalt principles in mind . . . to see if the design suggests any relationships between elements that you do not intend" (p. 23).

USE SUPPLEMENTAL CUES TO GUIDE USERS' ATTENTION

In addition to using design strategies to ensure that data symbols and patterns are easily perceptible, it may also be necessary to provide scaffolds that further encourage users to engage with a visualization's critical features, patterns, operators, and controls. Cueing user attention to critical elements can help minimize extraneous cognitive load by reducing the amount of searching or scanning necessary to locate areas or items of interest in a display (Dwyer, 1978; Jeung, Chandler, & Sweller, 1997; Mayer, 2005). Possible strategies include the following:

- Using automatically processed attributes, such as color, contrast, and motion, to highlight features or areas of a visualization
- Providing verbal information that directs users' attention to particular features or areas of a visualization
- Labeling critical features and patterns within a visualization
- Guiding interactions with the visualization
- Providing multiple modes of cueing to accommodate users with a range of expertise in a range of contexts

USE COLOR APPROPRIATELY

When applied strategically and appropriately, color can support users in effectively processing and interpreting data representations. We acknowledge that the exact colors that a user ultimately perceives are influenced by a number of factors (e.g., the lighting in the room, the monitor in use, etc.) (Ware, 2000), but there are a number of factors well within the interface designer's control that contribute to the effective use of color. It is important that an interface offer visualizations with color palettes that help users rapidly and accurately perceive the features and patterns represented.

First, consider color's strengths and weaknesses in the context of visual processing. While highly effective when used for labeling or categorization purposes, color hue is relatively poor at depicting an object's shape or form, fine details, motion, or depth. For these purposes, evidence shows that variations in luminance are more effective (Ware, 2000).

Second, interface designers must make colors perceivable in order for them to be useful. As mentioned earlier, colors that appear in the peripheral areas of the visual field are more difficult to perceive and distinguish. However, even in the center of the visual field where the retina's color-processing cones are plentiful, there are factors that can lead to inaccurate color perception as users attempt to distinguish *between* colors in visualizations featuring multiple hues. Color vision deficiencies, often referred to as color-blindness, affect approximately 8 percent of males and .5 percent of females. These deficiencies make it difficult or impossible for some users to distinguish between certain hues in normal lighting conditions. The most common form of color-blindness occurs in the red-green processing channel, resulting in difficulties distinguishing between red and green hues. The website Vischeck.com allows you to check webpages and images to see how they look to viewers with various color deficiencies (Dougherty & Wade, 2008). Even for those with "normal" vision, the perception of any individual color is influenced by its hue and the degree of similarity of surrounding colors (Light & Bartlein, 2004; Ware, 2000).

Finally, using colors that naturally “pop out” during initial visual processing, colors that are easy to remember, and palettes that are sensitive to cultural conventions can cue user attention to important features and patterns and facilitate accurate interpretation (Rheingans & Landreth, 1995; Ware, 2000).

The following are some specific strategies for using color appropriately in data interfaces:

- Use color hue to categorize or label visualization features.
- Avoid relying on color hue to indicate shape.
- Offer color schemes and palettes that address the needs of color-processing deficiencies.
- Use colors that are cross-culturally significant and recognizable: red, green, yellow, and blue.
- When using color to represent categories of data, choose colors that are as distinct from each other as possible. Red, green, yellow, blue, black, and white are the best choices. If more than six colors are needed, pink, cyan, gray, orange, brown, and purple are the next best choices.
- Use color in ways that are intuitive or familiar to the target audience.

REFERENCES

Brennan, N., & Lloyd, R. (1993). Searching for boundaries on maps: The cognitive process. *Cartography and Geographic Information Systems, 20*(4), 222-236.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education, 90*(6), 1073-1091.

Dougherty, R. F., & Wade, A. R. (2008). *Vischeck.com.* Retrieved from http://www.vischeck.com

Dwyer, F. M. (1978). *Strategies for improving visual learning: A handbook for the effective selection, design, and use of visualized materials.* State College, PA: Learning Services.

Germann, P. J., & Aram, R. J. (1996). Students' performances on the science processes of recording data, analysing data, drawing conclusions and providing evidence. *Journal of Research in Science Teaching, 33,* 773-798.

Jeung, H., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology, 17,* 329-343.

Johnson, J. (2010). *Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules.* Burlington, MA: Morgan Kaufmann.

Light, A., & Bartlein, P. J. (2004, October). The End of the Rainbow? Color Schemes for Improved Data Graphics. *EOS Transactions, 85,* 385-391.

Lloyd, R. (1997). Visual search processes used in map reading. *Cartographica, 34*(1), 11-32.

Lloyd, R., & Bunch, R. L. (2003). Technology and maplearning: Users, methods, and symbols. *Annals of the Association of American Geographers, 93,* 828-850.

Lloyd, R. E., Hodgson, M. E., & Stokes, A. (2002). Visual categorization with aerial photographs. *Annals of the Association of American Geographers, 92*(2), 241-266.

Lowe, R. K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction, 13,* 247-262.

Mathewson, J. H. (1999). Visual-spatial thinking: An aspect of science overlooked by educators. *Science Education, 83,* 33-54.

Mayer, R. E. (Ed.). (2005). Cambridge handbook of multimedia learning. New York, NY: Cambridge.

Rheingans, P., & Landreth, C. (1995). Perceptual Principles for Effective Visualizations. In G. Grinstein & H. Levkowits (Eds.), *Perceptual Issues in Visualization* (pp. 59-74). Berlin, Germany: Springer-Verlag.

Rieber, L. P. (1991). Effects of visual grouping strategies of computer-animated presentations on selective attention in science. *Educational Technology, Research, and Development, 39*, 5-15.

Ware, C. (2000). *Information Visualization: Perception for Design.* San Francisco, CA: Morgan Kaufmann.

Ware, C. (2008). *Visual Thinking for Design.* Burlington, MA: Morgan Kaufmann.

Wolfe, J. M., Klempen, N., & Dahlen, K. (2000). Post-attentive vision. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 693-716.

Cross-Cutting Guideline 3: Enable Customization

Rarely in learning—or in life, for that matter—does one size fit all. The need for customization (also referred to as "personalization of learning") is indisputable. Neurology, cognitive psychology, and education research make clear that heterogeneity is the name of the game. Every individual has "many different modules and distributed processes for learning . . . not just one or two generalized learning capacities" (Rose, Meyer, & Hitchcock, 2005, p. 23). Furthermore, individuals vary greatly from each other. To state it simply, different brains process information differently. Students will exhibit diverse aptitudes, skills, difficulties, and preferences. They will also bring to learning tasks variations in prior knowledge, based on their previous educational experiences, ideas they've drawn from everyday life, and schemata they have evolved to date.

And finally, we repeat a recurring mantra: Novice learners will differ from scientists who have amassed considerable scientific expertise—in both their modes of reasoning and their understanding of concepts and "big ideas," which influence what they notice and focus on; the ways they organize, represent, and interpret information; and their abilities to recall, reason, and problem-solve (Herr, 2007; Roschelle, 2012).

Heterogeneity necessitates a software architecture designed to afford flexibility, creating functionalities that allow learners to capitalize on strengths and preferences and, conversely, that help minimize the extraneous cognitive load (requirements irrelevant to the fundamental learning task at hand) and other hindrances that can limit learning. Enabling the customization of learning experiences is a central task for the developer, for as Rose et al. (2005) point out, "Barriers to learning occur in the interaction of the student with the curriculum—they are not inherent solely in the capacities of the learner" (p. 20). The customization guidelines below—related to Universal Design for Learning, user control, and supports for student learning—articulate means for reducing barriers to and facilitating learning.

INCORPORATE UNIVERSAL DESIGN FOR LEARNING

The Universal Design for Learning framework (Rose & Meyer, 2002; Rose et al., 2005) highlights the extensiveness of individual differences and the promise of flexible digital materials for optimizing learning given these differences (Dolan, Hall, Banerjee, Chun, & Strangman, 2005). Grounded in complex research on the neuroscience of learning, the three tenets of Universal Design for Learning may be summarized as follows:

- **Provide multiple means of representation.** Give students options for recognizing and understanding information and concepts via varied representations, including text, text-to-speech, concept maps, images, tables, graphs, simulations, and animations.
- **Create Support For Diverse Strategic Processes.** "Strategic processes" (branded "the how of learning" by the Center for Applied Special Technology http://www.cast.org/udl) are both the varied ways we express our ideas and the ways we plan, perform, and monitor our skills and actions. To provide multiple means for students to express their understanding, developers should create functionalities whereby students can show what they know, via, for example, recording, writing, illustrating, making a concept map, and/or creating maps and other visual representations.

  One way to facilitate students' development and execution of strategies and skills is to provide software utilities, such as tools for highlighting and tagging critical features in text, images, tables, maps, and graphs, and for creating tables and graphs. Another vital aid for strategies and skills is a digital "mentor" (aka "agent" or "reflective wizard"), which the student can activate in order to receive various types of scaffolding—for example, to identify critical features, big ideas, and relationships; to obtain background knowledge; to receive metacognitive help (i.e., support in monitoring one's learning and in knowing when and how to use particular strategies for learning); and to model "expert" performance. (For more on scaffolding, see Enable the Provision of Supports For Learning, below.)

- **Provide multiple means of engagement.** Being engaged and motivated is a sine qua non of learning. The developer can set the stage for engagement by providing students with user control (i.e., their choice of features to access), opportunities to interact with the software, scaffolding to attract and maintain student attention, and provision of the appropriate level of challenge.

The National Center on Universal Design for Learning offers the latest Universal Design for Learning guidelines for developers and educators. UDL Guidelines—Version 2.0 is available at http://www.udlcenter.org/aboutudl/udlguidelines.

AFFORD STUDENTS CONTROL

Congruent with the tenets of Universal Design for Learning are two recommendations for affording students choices as they interact with the material:

- **Allow users to control the pace at which information is presented.** Pacing provides the opportunity for students to adjust the rate of presentation of material to their individual cognitive needs. The research on pacing in digital environments focuses almost exclusively on animations and yields highly nuanced results. It is reasonable to recommend that if an animation is to be used, students should be able to pause and continue the animation as needed, so that "new information can be integrated into existing knowledge structures at a rate that reflects the capabilities and needs of the learner" (Plass, Homer, & Hayward, 2009, p. 48. It should be noted, though, that students do differ in how effectively they use interactivity (Keehner, Hegarty, Cohen, Khooshabeh, & Montello, 2008). (For a discussion of the complexities of using and pacing animations, see IV. Specific Considerations and Guidelines: Animations.)
- **Allow users to manipulate the parameters of representations.** When working with a map, for example, users should be able to change the scale, map projection, level of generalization, and color settings (Cartwright et al., 2001). Similarly, in working with graphs, students should be afforded the capacity to create graphs and to manipulate the graphs presented to them—for example, by labeling axes, highlighting a portion of a graph and labeling it, plotting data, and changing time spans of data (Tinker & Tinker, 2005).

ENABLE THE PROVISION OF SUPPORTS FOR LEARNING

Related to Universal Design for Learning and student control are certain features—scaffolding, expert guidance, monitoring support, and automating tools—that can facilitate learning. When developing such user-triggered help features, the software designer can choose among multiple formats, such as buttons, help icons, drop-down menus, point and click, pop-up boxes, agents, or avatars, as well as tools that perform various functions. Other considerations regarding supports for learning are as follows:

- **Incorporate the means for scaffolding.** While there are many definitions of the term "scaffold," it has traditionally been defined as a feature that supports the learner in accomplishing a task that is slightly beyond his or her current competence (i.e., a task that she or he could not do alone). This support complements and builds on the learner's existing abilities, allowing students to "participate at an ever-increasing level of competence" (Palincsar & Brown, 1984, p. 122) and "succeed in more complex tasks that would otherwise be too difficult, and . . . draw from that experience and improve in process skills and/or content understanding" (Reiser, 2004, p. 282).

  Scaffolds serve multiple purposes, ranging from reducing extraneous cognitive load to helping the student compensate for learning challenges, such as low spatial ability. The types of scaffolds are many, including advice, hints, cues, questions, sentence starters, explanations, concept maps, illustrations, tools, models, and glossaries. The use of such scaffolds is often "faded" (reduced or eliminated over time) as students develop competence.

- **Incorporate the capacity to provide "expert" guidance.** The knowledge and skills that experts bring to a task are greatly beyond a student's current competence. Providing the means for students to access expert guidance in the digital environment—via agents, avatars, and the other formats mentioned above—is critical. Types of expert guidance that can be offered include the following:
  - Cues for focusing on and recognizing important features and patterns
  - Examples (models), at a level of complexity appropriate for students, of how experts organize and solve particular kinds of problems
  - Access to relevant content knowledge, in the form of explanations or access to a glossary
  - Explanations of and rationales for expert practice
  - Reminders of criteria that students should apply to their work
- **Provide supports for students to monitor their learning.** In both science professions and the classroom, the practice of scientific inquiry is highly complex; as Quintana (2010) has opined, inquiry practices have "a wicked nature," given the complexities of the problem-solving processes involved, "where learners must constantly make several decisions to navigate this wide range of activities in a non-linear, iterative, opportunistic manner" (p. 1). Developers can provide software tools that encourage and help students keep track of, organize, and monitor their science practices and findings. As researchers have pointed out, software prompts can remind and/or provide encouragement about what steps to take (Davis & Linn, 2000; Reiser, 2004); graphical organizers or other notations can support students in considering possible actions relevant to each stage of the process and in planning and organizing their problem-solving (Quintana, Eng, Carra, Wu, & Soloway, 1999); and representations can help learners track and monitor the steps they have taken (Collins & Brown, 1988).
- **Create tools that automate the less essential aspects of learning tasks.** Software tools, such as calculators and graphing utilities, can minimize the cognitive load of certain aspects of tasks that are not central to successful learning, thereby allowing students to focus on the more conceptually important aspects of the learning task (Salomon, Perkins, & Globerson, 1991, cited in Reiser, 2004).

PROVIDE THE FUNCTIONALITY TO "PRESET" (CONSTRAIN) ACCESS TO INFORMATION AND TOOLS

Despite the advantages of customization, a caveat is in order: Providing students with virtually unlimited customization choices can be counterproductive (Fabrikant et al., 2008, cited in Nossum, 2010). For certain learning tasks, curriculum developers and teachers may deem it important to limit access only to relevant information, thus reducing the cognitive load of a particular task (Betrancourt, 2005). Quintana et al. (2004) caution designers to "incorporate . . . functional modes that offer only certain relevant subsets of software functionality" in order to constrain "what software tools are provided and when" (p. 363); imposing such constraints will help users avoid the confusion that may result when all possible tools and functions are available. The National Research Council (2006) has also recommended making software customizable, for example, by "making it possible for teachers to hide or expose functionality as needed" (p. 9) and by adopting an open system architecture.

There is no universal metric for balancing the dangers of overloading students against the advantages of providing them with sufficient choices to accommodate their particular needs. The answer seems to lie in malleability: designing an interface that offers as full a suite of options as possible along with the means to adjust availability at particular times for particular students in particular contexts.

REFERENCES

Betrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 287-296). Cambridge, England: Cambridge University Press.

Cartwright, W., Crampton, J., Gartner, G., Miller, S., Mitchell, K., Siekierska, E., & Wood, J. (2001). Geospatial Information Visualization User Interface Issues. *Cartography and Geographic Information Science, 28(*1), 1-19.

Collins, A., & Brown, J. (1988). The computer as a tool for learning through reflection. In H. Mandi & A. Lesgold (Eds.), *Learning issues for intelligent tutoring systems* (pp. 1-18). New York, NY: Springer-Verlag.

Davis, E. A., & Linn, M. C. (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. *International Journal of Science Education, 22,* 819-837.

Dolan, R. P., Hall, T. E., Banerjee, M., Chun, E., & Strangman, N. (2005, February). Applying Principles of Universal Design to Test Delivery: The Effect of Computer-based Read-aloud on Test Performance of High School Students with Learning Disabilities. *The Journal of Technology, Learning, and Assessment, 3*(7).

Herr, N. (2007). How experts differ from novices. *Science Teaching Series.* Northridge. CA: California State University, Northridge. Retrieved from http://www.csun.edu/science/ref/reasoning/how-students-learn/2.html

Keehner, M., Hegarty, M., Cohen, C., Khooshabeh, P., & Montello, D. (2008). Spatial reasoning with external visualizations: What matters is what you see, not whether you interact. *Cognitive Science, 32,* 1099-1132.

National Research Council (2006). *Learning to think spatially.* Washington, DC: The National Academies Press.

Nossum, A. (2010, September 15-17). *Semistatic animations—Integrating past and future in map animations.* Presented at GIScience 2010, Zurich, Switzerland.

Palincsar, A. M., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 2, 1*17-175.

Plass, J. L., Homer, B. D., & Hayward, E. O. (2009). Design factors for educationally effective animations and simulations. *Journal of Computing in Higher Education,* 21(1), 31-61.

Quintana, C. (2010). *Invited talk: Making the implicit explicit: Issues and approaches for scaffolding metacognitive activity. Cognitive and Metaconitive Educational Systems: Papers from the AAAI Fall Symposium.* Center for Highly Interactive Classrooms, Curricula, and Computing, School of Education, University of Michigan, Ann Arbor.

Quintana, C., Eng, J., Carra, A., Wu, H., & Soloway, E. (1999). Symphony: A Case Study in Extending Learner-Centered Design Through Process-Space Analysis. *In Human Factors in Computing Systems: CHI '99 Conference Proceedings.* Pittsburgh, PA: ACM Press.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fetz, E., Duncan, R. G., Kyza, E., Edelson, D., & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences, 13*(3), 337-386.

Reiser, B. J. (2004). Scaffolding Complex Learning: The Mechanisms of Structuring and Problematizing Student Work. *The Journal of the Learning Sciences, 13*(3), 273-304.

Roschelle, J. (2012). *Learning in Interactive Environments: Prior Knowledge and New Experience.* Dartmouth, MA: Institute for Inquiry, University of Massachusetts, Dartmouth. Retrieved from http://www.exploratorium.edu/IFI/resources/museumeducation/priorknowledge.html

Rose, D. H., & Meyer, A. (2002). *Teaching every student in the digital age: Universal Design for Learning.* Alexandria, VA: ASCD.

Rose, D., Meyer, A., & Hitchcock, C. (2005). *The universally designed classroom: Accessible curriculum and digital technologies.* Cambridge, MA: Harvard Education Press.

Tinker, R., & Tinker, B. (2005). *Flexible software design: Software to accommodate individual learning differences.* Concord, MA: Concord Consortium. Retrieved from http://www.concord.org/sites/default/files/pdf/tinker-fsd-2005.pdf

Specific Considerations and Guidelines

IV. SPECIFIC CONSIDERATIONS AND GUIDELINES

IV. SPECIFIC CONSIDERATIONS AND GUIDELINES

Accessing Data

This section focuses on the first thing that students will encounter when using an online data portal: the user interface that allows them to select and view data.

At the time of this writing, very little research has been directly targeted at how to design Web interfaces that facilitate novice access to professional scientific databases. However, if this first step in the process of working with data isn't handled well, students (and most pre-college teachers) will go no further. There is so much intrinsic difficulty associated with working with large, diverse scientific data sets—getting the usability right is even more critical than it typically would be.

The considerations and guidelines in this section are based on the most relevant research we were able to find, as well as the experience of the project team and our advisors. It should be emphasized that beyond the suggestions in this section that specifically apply to data portals, it is especially important that developers adhere to good general design principles for user interfaces and universal accessibility (see Johnson, 2010; National Center on Accessible Instructional Materials, n.d.; Shneiderman & Plaisant, 2010). (For emphasis, this is also repeated as GUIDELINE 8.)

The guidelines we present here focus on factors that we believe are critical to involving science classes in the use of online scientific databases. Of course, it's very important that any computer interface be thoroughly tested with the intended users—which means going beyond the small group of highly motivated teachers who are willing to deal with a difficult interface.

CONSIDERATIONS

WHAT ARE WE TALKING ABOUT?

Software and curriculum developers who tackle the task of providing classrooms with access to online databases could take several different approaches, for example:

- Trying to build an education interface "on top of" an expert interface, providing definitions and other supports to scaffold students' experiences
- Writing a curriculum that explains the steps that students should follow to interact with a professional interface
- Building a brand-new interface focused on student users

It is the strong view of the authors and the members of the Advisory Board that the last approach, building a new interface, is the preferred one. Interfaces intended for expert users are simply too laden with expert terminology that needs to be defined, and with data exploration processes and displays that are unfamiliar and non-intuitive to novice users.

At times, building on top of an expert interface or simply relying on curriculum or teacher scaffolding is the only course available. However, large scientific cyberinfrastructures are currently being built with the intention that the data they provide will be public data from the outset; anyone interested will be able to get the data and customize the data stream to create his or her own "laboratory."

The considerations and guidelines presented in this section focus on the graphical user interface itself—the website that students and teachers encounter and interact with when they want to access and explore data. We envision what this Web interface should look like and how it should work (though we leave the nuts and bolts of how it gets that way to others!).

A data access portal intended for student users should have a simple goal – it should make it easy for students to see the data that are available, and to explore, think about, and generate questions about these data.

WHAT SHOULD THE INTERFACE DESIGNER CONSIDER?

Novices' lack of familiarity with domain vocabulary, and with the technologies associated with remotely collected data, can frustrate their ability to find appropriate data. The typical Web interface to a large, complex scientific database poses significant obstacles to students. For example, a student looking for an answer to the question "What was the ocean temperature in the eastern Pacific last month?" might confront a page such as that shown in Figure 5—an interface designed for expert users.

The example in Figure 5 reflects the fact that large scientific databases can offer dozens of data parameters (as well as multiple data sets for a particular parameter, such as ocean temperature), collected using a wide variety of instruments or representing the output of computer models. Students may be confronted with acronyms, terms, and units that are unfamiliar to them (Edelson, Gordin, & Pea, 1997). For example, data sets may be grouped by the names of the research programs they are associated with, which will have meaning only for research scientists. Users may be prompted to use key-word searches if they can't figure out which data option to click on, but because there is often a fundamental mismatch between the students' vocabulary and the system vocabulary, the key-word search option is not helpful (Butcher, Bhushan, & Sumner, 2006). Curriculum developers and teachers interested in helping students use these data portals must resort to tutorials or instruction sheets with long lists of steps that, because they carry no meaning for students, are tedious and distract from students' learning.

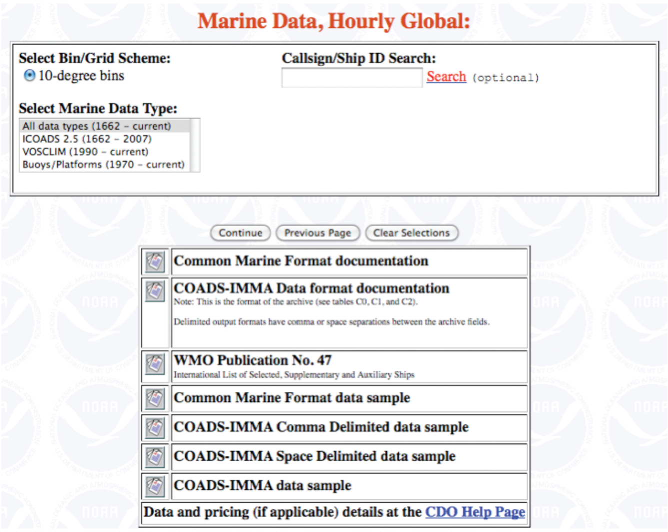


Figure 5. The sheer number of data parameters offered by large scientific databases and the terminology used to describe them can be overwhelming to users, as illustrated by the data selection interface shown above. (Source: NNDC Climate Data Online, NOAA Satellite and Information Service. Retrieved from http://www7.ncdc.noaa.gov/CDO/CDOMarineSelect.jsp.)

If a digital information database does not support easy and intuitive browsing and discovery, this can pose a significant barrier to the use of the data. Johnson (2010) opens his book *Designing with the Mind in Mind* with this thought:

> *We perceive, to a large extent, what we expect to perceive. Our expectations—and therefore our perceptions—are biased by three factors:*
>
> - *The past: our experience*
> - *The present: the current context*
> - *The future: our goals (p. 1)*

The experience, knowledge, context, and goals that students bring with them when they interact with a data portal will be very different from those of expert scientists (or the programmers responsible for creating the interface). These differences mean that students will have different expectations of a data interface than an expert user will. Research on the design of user interfaces has shown that users of computer software and websites often click on buttons without looking at them carefully (Johnson, 2010), and their instincts will be incorrect if their expectations don't match what is on the screen. Medyckj-Scott and Blades (1992) discuss accessibility issues associated with spatial information systems, such as Geographic Information Systems (GIS), and argue that if a user's schema for interacting with the system is inaccurate or incomplete, the user will expect a GIS to behave in a way that doesn't match the system's actual behavior. Assuming that the user doesn't get frustrated and give up, he or she is likely to restrict interactions with the system to "safe routes" and fail to explore and use all the features of the GIS.

GUIDELINES

ACCESSING AND VIEWING DATA SHOULD BE FAST AND EASY (GUIDELINES 1–8)

The goal of the interface should be to maximize the degree to which students are able to use what they already know (their existing schemata) and their automatic cognitive processes to quickly find and view data, focusing their attention on important learning tasks. Beyond simply searching data sets, students should be supported in exploring the data and developing their own questions about them. This is the essence of scientific work.

Guideline 1: There should be low or no barriers to downloading and visualizing a data set.

Students should be able to get very quickly to the point where they are able to explore and generate questions about the data. One possible approach is to provide interesting default data sets that are immediately available, which are in turn connected to a larger variety of data that students can access on demand (see also Guideline 8). Downloading data should connect seamlessly with generating data visualizations, and students should be able to modify visualizations or change data sets without going back to step one. The default settings for visualizations should be customized and optimized to the data so that the visualizations are immediately viewable, minimizing the need to go through steps to change settings. For example, the interface currently under development as part of the NSF-funded Analyzing Ocean Tracks Project provides students with immediate access to visualizations of elephant seal tracking data, where default settings such as map scale and track color are optimized for the display students initially encounter (see Figure 7).

Guideline 2: Make the important controls stand out

> *Interactive systems should be designed so that the scent is strong and really does lead users to their goals.*
> *—Jeff Johnson, Designing With the Mind in Mind, 2010 (p. 100)*

Harrower and Brewer (2011) point out that learning how to use an interface involves "at least two critical steps: knowing what the buttons do and knowing the order in which to use them" (p. 266). They lament the fact that

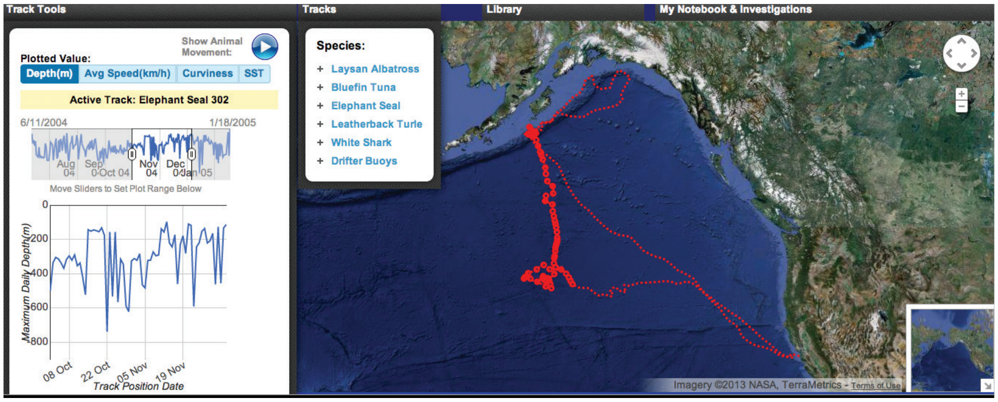


Figure 7. The landing page for this interface to marine animal tracking data gives students immediate access to elephant seal tracking data in both georeferenced and graph visualizations. The default settings are optimized so that the track and orienting land masses are clearly visible. (Source: Analyzing Ocean Tracks Project. Retrieved from http://oceantracks.org).

software engineers often hide key controls or options deep within the interface, rather than organizing them so that the most important controls are highest in the visual hierarchy. Making the most important controls stand out means keeping the webpages simple and clear of distracting information, such as advertisements, extraneous links, images, and animations (Nivala, Brewster, & Sarjakoski, 2011). Harrower and Brewer (2011) also suggest having controls appear only when they are needed to guide the user through the interface. Medycki-Scott and Blades (1992) suggest that presenting a map of the system (see also Guideline 4) and/or signposts to show where to go next can reduce the possibility that users will become lost. Gordin, Polman, and Pea (1994) suggest using a standard and familiar English-style text-informed arrangement for information on a webpage—arranged from general to specific, from controls to results, from left to right, and from top to bottom. Clearly marking steps that have been completed (by, for example, changing the color of a button) also helps users stay on track (Johnson, 2010).

Guideline 3: Minimize expert terminology

Avoid using expert terminology or jargon (including that associated with the underlying programming architecture) and use familiar, task-focused descriptors on the user interface (Edelson, Gordin, & Pea, 1997; Johnson, 2010). Research has found that skilled (fast) reading is an automatic process of recognition and uses a different part of the brain from unskilled reading, which involves sounding out words and deciphering their meaning (Johnson, 2010). Thus, unfamiliar vocabulary can significantly disrupt students' thinking processes and distract their attention from more important learning tasks.

Guideline 4: Simplify and structure information

When navigating through a website, it is easier for people to scan quickly and automatically for relevant information when that information is presented in a concise, structured way (Johnson, 2010). For example:

- **Visually group related information** (such as labels and values) or controls by proximity, without adding bordered boxes or other labels that add to the visual clutter.
- **Have new information pop up over related content** to shift students' focus temporarily without losing their connection with larger goals (Johnson, 2010).
- **Provide options that students can choose,** rather than requiring them to type in search terms—recognition is much easier for humans than recall.
- **Use pictures** as a clear and concise way to convey function, particularly if their meaning is consistent across the interface and, ideally, similar interfaces (Johnson, 2010).
- **Minimize text on webpages** and use restricted and highly consistent vocabulary that is understandable to a broad audience (see also Guideline 3).
- **Use text fonts that are easy to read**—sans serif for shorter bits of text in a screen display and serif for longer narrative reading (Evergreen, 2012). Avoid busy backgrounds and colors that interfere with each other.
- **Arrange text using a visual hierarchy,** with headings, bullets, and tables that make the text easy to scan and read automatically (Johnson, 2010).

Guideline 5: Bridge to the familiar

Today's students are accustomed to poking around in a computer interface and figuring things out for themselves. However, there are many wrong and potentially confusing paths to go down in an expert scientific data interface. This not only befuddles and frustrates students, it also diverts cognitive resources from the important learning tasks. Therefore, it's critical to design the interface so that it bridges to the familiar by using layouts, tools, and features that students are likely to have encountered before. For example, students are likely to have used certain common types of map tools for selecting areas of interest by clicking on an icon and drawing a box or using a sticky hand to grab and move to another location. Such features as drop-down menus and mouse-over pop-ups are also likely to be familiar. Since it would be an overwhelmingly impossible task to standardize the design of these features for the interfaces students encounter outside of the classroom, educational interface designers must be aware of and employ what is commonly used in public Web interfaces.

Medyckyj-Scott and Blades (1992) highlight the usefulness of metaphors in helping novices develop schemata for how a Web interface works. An example is the desktop metaphor commonly used by many computer systems. Cartwright (2011) suggests a number of metaphors that could be applied, including a guide metaphor that "navigates for the user and takes the user to places that the guide computes as necessary" (p. 180) to accomplish the intended task.

Because Web interfaces to scientific databases have unique functional requirements and intended uses, they can't simply apply features and metaphors from nonscientific websites that are in broad public use. Therefore, students and teachers will need to invest time in learning how to use these interfaces. This makes it important to move toward standardizing the way that education interfaces to scientific data portals work, so that classroom users don't have to relearn the system each time they encounter a new data interface. (See also Guideline 7.)

Guideline 6: Make it easy to download the data for use with other tools

Designing a data interface that will work for all education users for all purposes is, of course, an impossible endeavor. Therefore, there should be an easy way for users to download the data so that they can be analyzed using customized tools that are developed by others.

Guideline 7: Strive for both internal and external consistency

Students' interactions with any computer interface will be influenced by what they've encountered before and what they can easily and quickly learn through interaction with the new interface. It is important to note that if interfaces to scientific databases are not standardized, and if students or teachers have to learn a new system of interaction each time they encounter a new database, this will significantly hinder the system's usefulness, particularly in today's time-crunched classroom (Cartwright et al., 2001). Research is needed to determine the degree of standardization that is possible, particularly when a common interface will be used for diverse data sets and diverse purposes. However, there is evidence that a common interface can work in different contexts. For example, the Lamont Data Viewer, which was originally developed for research, has been modified for educational use and successfully used by students taking courses in a variety of disciplines (Hays, Pfirman, Blumenthal, Kastens, & Menke, 2000).

Internal consistency is also critical to expedite students' learning of how to interact with a particular data access interface. Information and controls should be placed in consistent locations and serve consistent functions. They should also look the same, so they will be noticed and recognized quickly (Johnson, 2010; Shneiderman & Plaisant, 2010).

Guideline 8: Follow standard rules for good usability

Researchers have been learning about how to build good user interfaces for several decades, as this is of critical importance to a much broader audience than just those developing interfaces to scientific databases. Thus, a number of excellent resources are available that interface designers should be familiar with to ensure the usability of interactive websites, particularly those aimed at non-expert audiences. Books such as *Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules* by Jeff Johnson (2010) and *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, by Shneiderman, Plaisant, Cohen and Jacobs (2010) provide design guidelines and many excellent examples, as well as the underlying research.

It should be also be emphasized that interfaces intended for classroom use must be universally accessible. The National Center on Accessible Instructional Materials (http://aim.cast.org) has compiled a great deal of information on this topic, including a long list of Web resources. The DIAGRAM project (http://diagramcenter.org) aims to standardize how various alternative representations for complex graphics can be organized and made available in a way that can be used by accessibility technologies (B. Goldowsky, personal communication, September 6, 2012).

USE ORGANIZATIONAL STRUCTURES THAT QUICKLY ENGAGE AND ALSO SUPPORT DEEPER EXPLORATION (GUIDELINES 9–12)

Guideline 9: Use hierarchical structures

A typical scientific data interface offers a huge variety of data sets for scientific use. To make these scientific databases accessible to students, it is important to initially limit the number of choices in educational interfaces, then work up to progressively deeper engagement with a variety of data sets as students' interest and sophistication develops. Hierarchical structures are an intuitive way to organize information so that it doesn't initially overwhelm students, but allows more in-depth exploration when desired.

Decisions about which data sets to offer initially should be based on what is likely to be most understandable and interesting to students and most relevant to their learning tasks. One way to decide is to consider the topics that students are likely to care about and the questions they are likely to ask (D. Edelson, personal communication, January 11, 2011). This requires working with curriculum developers who know what will engage students and what content standards will motivate their curriculum. Edelson et al. (1997) also suggest providing data that represent familiar quantities and relate to human activity.

It should be emphasized that students may not know what data they want to view when they go to a data portal, so the initial choices should give them understandable and interesting options that will entice them to dig deeper. The organization should also make it easy for students to explore the variety of data that are available.

Fabricant (2000) emphasized the importance of presenting information spatially to take advantage of the strengths of human cognition, and also suggested a hierarchy for exploratory information-seeking based on the oft-repeated mantra by Ben Shneiderman (1998):

> *Overview first, zoom and filter, then details-on-demand (p. 523).*

In our case, this means first giving an overview of the entire data collection, then allowing users to zoom in on items of interest, filter out uninteresting items, and get details about a data set on demand.

Guideline 10: Use visual organizers to support conceptual browsing and discovery

To help students find data that are of interest to them or that they need to answer their questions, research suggests the use of visual conceptual organizers (Quintana et al., 2004), such as conceptual diagrams or representations of visual scenes through which users can access functionality. In the example in Figure 8, an energy balance diagram is used to help students first think about the factors they should investigate and then access the relevant data (as recommended by Edelson, Gordin, & Pea, 1999).

Simple concept maps (diagrams that show the relationships between concepts) can also be used to organize access to data. Butcher et al. (2006) researched the cognitive processes that students engaged in when working with concept maps (referred to as "strand maps") versus text search interfaces to digital libraries, and found that students were more likely to engage with germane science content during their search when they used the concept map interfaces.

Guideline 11: Allow for easy access to necessary background knowledge about the nature, source, and appropriate use of the data

To support students' selection of appropriate data, links to background information—including definitions of expert terms and unfamiliar units of measurement—should be provided. Given the challenge of engaging students with remotely collected data, it is particularly important to include descriptions of how the data were collected.

Guideline 12: Support sustained engagement with data

Education interfaces to professional data sets should encourage and support extended engagement with the data. There should be no school-specific barriers to continuing work with a particular data set and visualizations at home. This also will allow students who fall prey to the "addictiveness" of working with data to continue their explorations beyond the context of the classroom.

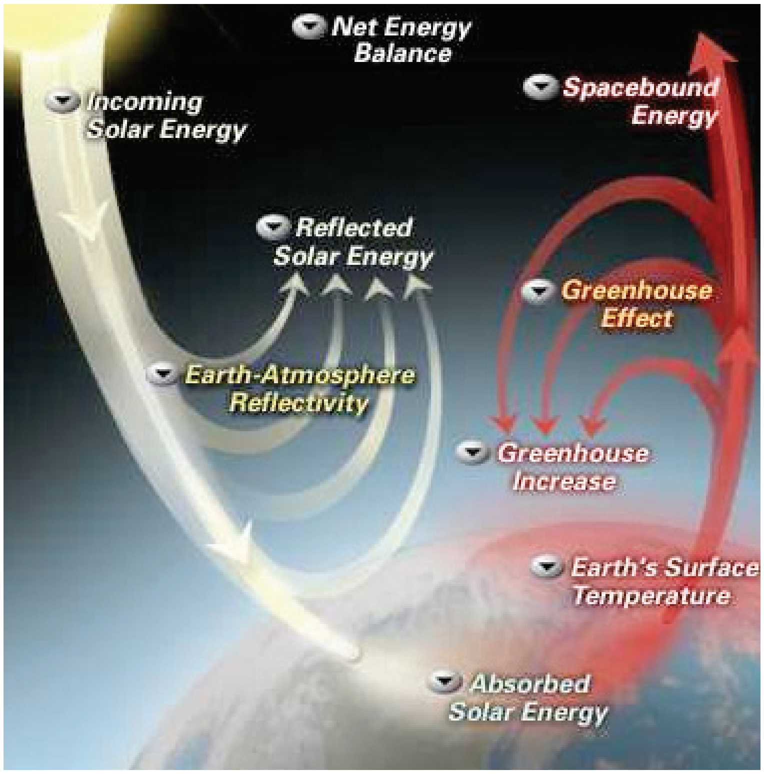


Figure 8. The interface to energy balance data in WorldWatcher, a geographic data visualization and analysis tool that provides users with diagrammatic interfaces for accessing data. (Source: GEODE [Geographic Data in Education] Initiative, Northwestern University. Retrieved from http://www.worldwatcher.northwestern.edu/softwareoverview-interface.htm.)

REFERENCES

Butcher, K. R., Bhushan, S., & Sumner, T. (2006). Multimedia displays for conceptual discovery: Information seeking with strand maps. *Multimedia Systems, 11*(3), 236-248.

Cartwright, W. (2011). Extending the map metaphor using Web delivered multimedia. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader* (chapter 2.11). Chichester, UK: Wiley-Blackwell.

Cartwright, W., Crampton, J., Gartner, G., Miller, S., Mitchell, K., Siekierska, E., & Wood, J. (2001). Geospatial information visualization user interface issues. *Cartography and Geographic Information Science, 28*(1), 45-60.

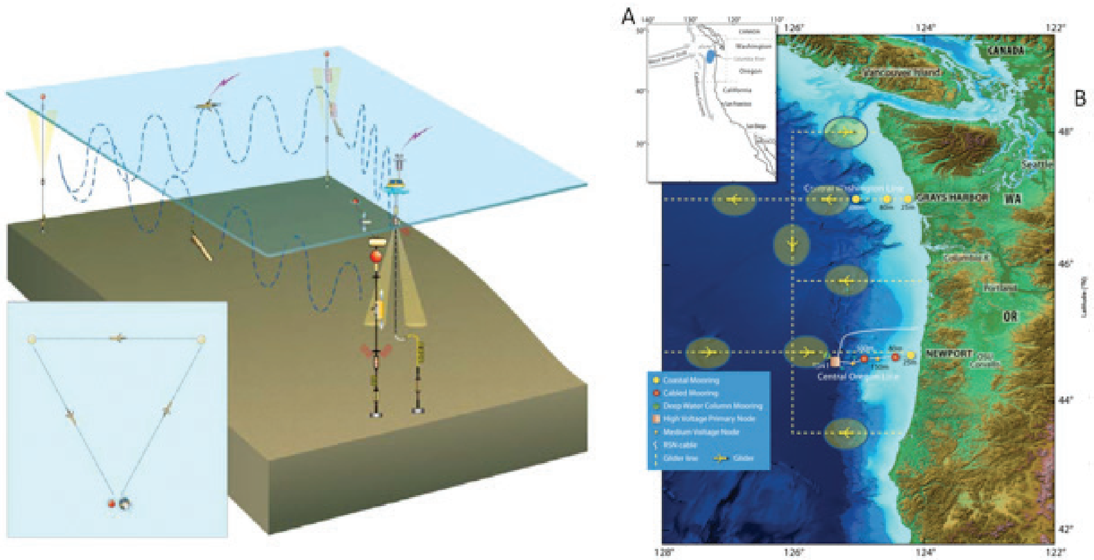


Figure 9. Left: The path of a robotic glider is charted off the coast of Washington and Oregon. The glider repeatedly dives and surfaces as it travels between buoys, collecting physical oceanographic measurements, such as temperature, salinity, and dissolved oxygen. (Source: Illustration by Jack Cook, copyright Woods Hole Oceanographic Institution) Right: Proposed Endurance Array mooring locations and glider lines off the coast of Oregon and Washington. (Source: Illustration by David Reinert, copyright Oregon State University)

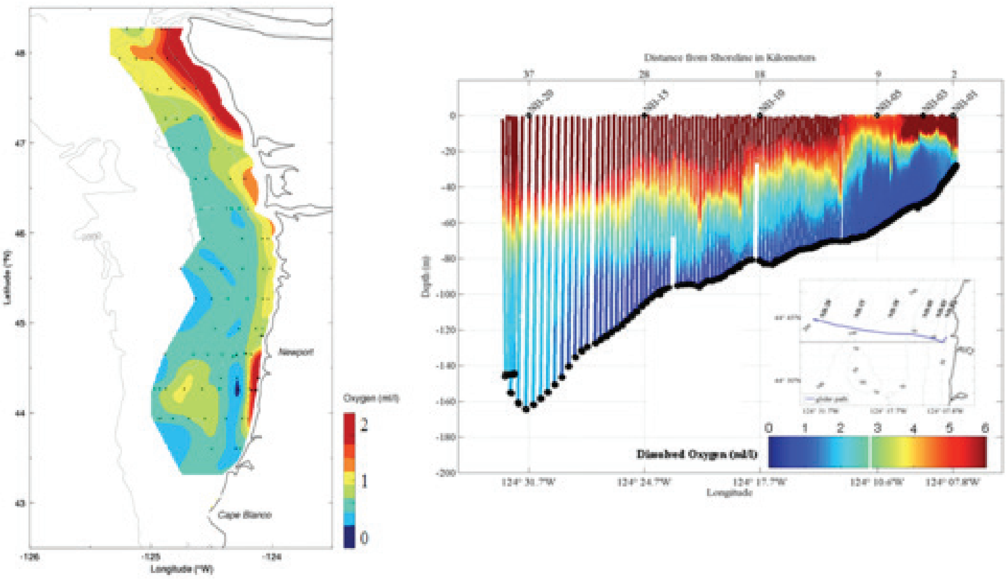


Figure 10. Left: This display shows the September 2007 distribution and extent of hypoxic waters along the Oregon-Washington shelf. (Source: Committee on Environment and Natural Resources. 2010. Scientific Assessment of Hypoxia in U.S. Coastal Waters. Interagency Working Group on Harmful Algal Blooms, Hypoxia, and Human Health of the Joint Subcommittee on Ocean Science and Technology. Washington, DC. Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/hypoxia-report.pdf.) Right: A cross-section showing dissolved oxygen concentrations. (Source: J. Barth, R. K. Shearman and A. Erofeev, Oregon State University, unpublished data)

Edelson, D., Gordin, D., & Pea, R. D. (1997, March 20-24). *Creating science learning tools from experts' investigation tools: A design framework.* Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Oak Brook, IL.

Edelson, D., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design, *The Journal of the Learning Sciences, 8*(3), 391-450.

Evergreen, S. (2012, September 25). *Data Visualization and Reporting.* Presentation given at EDC, Waltham, MA.

Fabricant, S. I. (2000). Spatialized browsing in large data archives. *Transactions in GIS, 4*(1), 65-78.

Gordin, D. N., Polman, J. L., & Pea, R. D. (1994). The Climate Visualizer: Sense-Making Through Scientific Visualization. *Journal of Science Education and Technology, 3*(4), 203-226.

Harrower, M., & Brewer, C. A. (2011). ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader* (chapter 3.8). Chichester, UK: Wiley-Blackwell.

Hays, J. D., Pfirman, S., Blumenthal, B., Kastens, K., & Menke, W. (2000). Earth science instruction with digital data. *Computers & Geosciences, 26*, 657-668.

Johnson, J. (2010). *Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules.* Burlington, MA: Morgan Kaufmann.

Medyckyj-Scott, D. J., & Blades, M. (1992). Human spatial cognition: its relevance to the design and use of spatial information systems. *Geoforum, 23*(2), 215-226.

National Center on Accessible Instructional Materials. (n.d.). *Get Started: All About AIM.* Retrieved from http://aim.cast.org

Nivala, A. -M., Brewster, S., & Sarjakoski, L. T. (2011). Usability Evaluation of Web Mapping Sites. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader* (chapter 4.11). Chichester, UK: Wiley-Blackwell.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fetz, E., Duncan, R. G., Kyza, E., Edelson, D., & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences, 13*(3), 337-386.

Shneiderman, B. (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (3rd ed.). Reading, MA: Addison-Wesley.

Shneiderman, B., & Plaisant, C. (2010). *Designing the User Interface: Strategies for Effective Human-Computer Interaction (5th ed.). Reading*, MA: Addison-Wesley.

Geo-Referenced Data Representations (Plan Views, Cross-Sectional Views, and 3D Views)

CONSIDERATIONS

WHAT ARE WE TALKING ABOUT?

Geographically referenced data sets—such as those associated with oceanographic, atmospheric, and geophysical databases—capture information about the geospatial distribution of measurements. Maps (plan views), cross-sections (depth views), and 3D visualizations (which present multiple spatial perspectives in one image) are powerful tools used by scientists to identify spatial patterns in these data.

These 3D data sets are collected throughout the ocean, atmosphere, and geosphere using gliders (autonomous robotic vehicles that travel between buoys in the ocean), combinations of satellites and ground stations or buoys, and seismic arrays of seismometers. An example of how a 3D data set is collected is provided in Figure 9.

Three-dimensional data can be displayed as plan views, or "views from above," that show the aerial distribution of measurements. These plan views may depict the aerial distribution of a particular type of data at the surface of the ground or ocean, or at a designated depth, elevation, or stratum. In many cases, depending on the purpose of the data analyses, the measurement may be averaged over a range of depths or elevations. For example, the map on the left in Figure 10 shows dissolved oxygen measurements collected by gliders, taken at the deepest points of the glider's path (presumed to correspond to a level just above the ocean floor, which varies in depth). Since the actual glider measurements occurred along transects between buoys, the map shows interpolated data between the actual measurements and represents the data in a color contour plot, with red showing areas with the highest concentrations of dissolved oxygen, and dark blue showing areas with the lowest concentrations.

Cross-sections show vertical "slices" of data along lines or transects. The cross-section on the right in Figure 10, for example, shows dissolved oxygen concentrations taken by the glider along one particular transect as it moved between buoys. Data for portions of the cross-section between the actual glider path are interpolated using computer models. Again, red shows the portions of the profile with the highest concentrations of dissolved oxygen and dark blue the portions with the lowest concentrations. However, note that the scale is different from the map on the left, so the colors have different meanings.

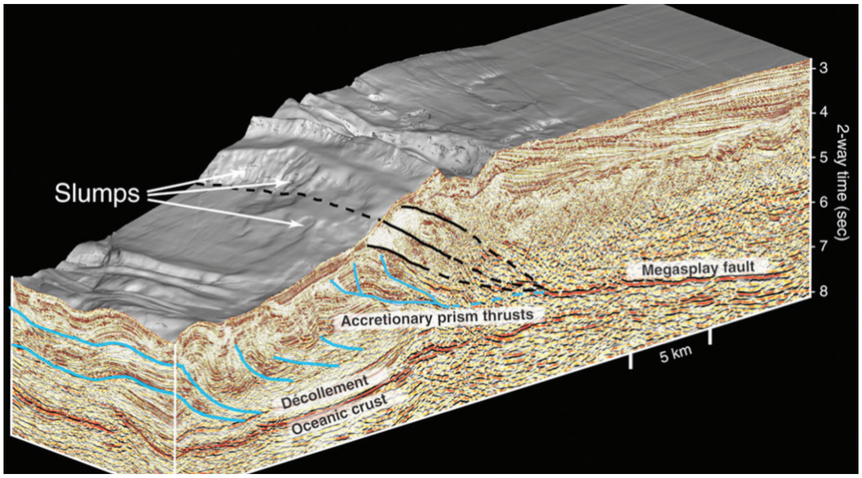


Figure 11. This three-dimensional block diagram is used to visualize data regarding the internal layers, fault, and slump structures within Earth's crust in a region of the Pacific seafloor. (Source: G.F. Moore, N.L. Bangs, A. Taira, S. Kuramoto, E. Pangborn and H.J. Tobin, 2007. Three-dimensional splay fault geometry and implications for tsunami generation. Science, 318, pp. 1128-1131.)

Various types of visualizations are also used to combine both horizontal and vertical components of three-dimensional data, such as the 3D block diagram in Figure 11.

Most of the relevant literature available at the time of this writing pertains to maps. More research is needed to understand how to help students understand cross-sectional views, which are commonly used in oceanography and the geosciences but much more rarely encountered elsewhere. Three-dimensional depictions and interactive visualizations show promise for helping students relate different data views. However, many of the tools that currently exist (such as the tool that allows users of Google Maps to draw a profile line on a map and generate a profile view) focus on conveying the shape of surfaces, not the three-dimensional distribution of subsurface measurements.

WHAT SHOULD THE INTERFACE DESIGNER CONSIDER?

Working with and thinking about geo-referenced data requires spatial and visualization abilities that vary significantly from one individual to another. Research indicates that each of us employ a different "cognitive style" or mode of information-processing, which is determined both by our innate ability and by our prior experience (Bunch & Lloyd, 2006). Some researchers classify individuals into three categories: verbalizers, object visualizers, and spatial visualizers (Chabris et al., 2006; Kozhevnikov, Kosslyn, & Shephard, 2005).

Although work in many professions primarily involves one type of visualization strategy, both *object visualization*—the ability to recognize patterns and shapes in a complex background and to distinguish among similar shapes—and *spatial visualization*—the ability to mentally rotate three-dimensional objects and to imagine and coordinate views from different perspectives—seem to be important for working with geo-referenced data (Ishikawa & Kastens, 2005; Kastens, 2010). As students view visualizations of these data, they must perceive the data accurately to make meaning of them and construct schemata. (These types of schemata may also be referred to as "spatial images" or "cognitive maps" [see, for example, Bednarz, 2001; Medyckj-Scott & Blades, 1992].) Accurate perception includes identifying the critical information in a complex image—for example, identifying the significant spatial patterns in a map (object visualization) and also identifying and coordinating those features in different data views, such as plan views and profile views (spatial visualization). As discussed in II. KEY UNDERPINNINGS: VISUAL PERCEPTION AND PROCESSING and SCHEMATA, the construction of robust schemata with respect to geo-referenced data displays is important because it allows students to build their ability to deal with increasingly complex data visualizations. These schemata preserve visual patterns and spatial relationships in our long-term memory so that they can be applied in single chunks in our working memory, increasing the amount of information that can be dealt with at a single time.

And here's the challenge: Research has provided evidence that people with high object visualization ability tend to have low spatial visualization ability, and vice versa (Kozhevnikov et al., 2005). In addition, these visualization abilities do not correlate with verbal ability. Students who are used to doing well because they have strong verbal abilities (which are emphasized in K-12 schooling) may struggle with such tasks as reading and working with maps that draw on their visual and spatial reasoning skills. At the same time, students who typically struggle with verbal tasks may excel at visual and spatial tasks (Ishikawa & Kastens, 2005). Although research indicates that spatial thinking is a skill that can be learned, it is not systematically taught in the K-12 curriculum (National Research Council [NRC], 2006). Perhaps because of this, even students at the college level tend to have poor understanding of geospatial concepts (NRC, 2006). All of this means that the interpretation of various geo-referenced data representations may be difficult for most students, but in different ways, depending on their particular cognitive abilities and previous experience.

The visual features and patterns that a novice student perceives in a geo-referenced data representation will not be the same as those perceived by an expert. As discussed in VISUAL PERCEPTION AND PROCESSING, our brains are selective about the information from our environment that ends up in a mental image, because the amount of new information that can be processed by our working memory is limited. As we build expertise, we develop schemata stored in our long-term memory that can be applied automatically, without making demands on our working memory. This allows us to more easily process more sophisticated data visualizations. But because novices don't yet have these expert schemata, their brains draw more heavily on the inherent evolution-based biases in our visual perception system. They will focus on features that naturally "pop out," such as yellow bands in a rainbow color palette. The result is that novices will not "see" the same image that an expert sees.

Geo-referenced representations of oceanographic and other Earth science data are likely to impose a high level of intrinsic cognitive load.

> *The maps are really nice to look at but they make no sense to me. I get [that] the colors mean things but I need to see a real graph. Maybe if we had time I would get used to the maps. I would like to, they probably tell more.*
> *—High school student participating in pilot-testing of lessons using maps of atmospheric data*

Scientific data visualizations that show spatial variations in measurements, such as temperature, dissolved oxygen, salinity, or chlorophyll at a particular time, are likely to be complex—both the patterns in these measurements and the way they are depicted will be unfamiliar and non-intuitive to many or most novice students. In addition to lacking prior knowledge of remotely collected data parameters, how they are measured, and what their distribution means, students will simply have difficulty reading a map or cross-section. For example, although they have ample experience recognizing the outline of continents on a world map or certain political boundaries, maps focusing more closely on regions with physical features alone are much less likely to be familiar. The geographic area and scale covered may not be clear to students, and they are also likely to be inexperienced in reading maps or cross-sections that use abstract elements, such as color and contours, to show different concentrations. Svenson and Kastens (2011) found that students had a great deal of difficulty understanding even a colored, shaded-relief global elevation map.

In addition, Earth systems phenomena involve a large number of interacting elements, and no one map or cross-section alone is likely to explain the patterns revealed therein. Therefore, multiple data representations must be related to each other, and their potential causal relationships should be considered, which will significantly increase the intrinsic (and hopefully germane) cognitive load.

Research has shown that high school students also have difficulty visualizing the hidden internal structure within 3D block diagrams that depict geo-referenced data (Kali & Orion, 1996). To maximize the amount of working memory available to deal with cognitively demanding tasks, it is critical to make design choices that reduce the extraneous cognitive load associated with creating, reading, exploring, and comparing geo-referenced data visualizations.

The creation of new geo-referenced data representations on a Web interface is likely to be particularly challenging for students. Online interfaces to scientific databases bring the capability for students to create their own unique data visualizations. Discovering patterns in data that perhaps no one has seen before is one of the most satisfying aspects of scientific work, and online databases have great potential to give students these types of experiences. Tools for creating data visualizations are currently available on many websites, bringing this capability to a large audience. However, students have very limited experience making even the simplest kind of map from direct observations (Svenson & Kastens, 2011), and maps generated from remotely collected or real-time data may be particularly confusing. Students are used to looking at textbook visualizations that are optimized to show certain features—whereas on a map these features may not stand out without adjusting the particular display, which will be a significant challenge for students to do well. As Perkins, Dodge, and Kitchin (2011) noted, the proliferation of online mapping tools "has led to a concern amongst many cartographers that we are entering an age of poorly designed, DIY [do it yourself] maps" (p. 197). Rogowitz and Treinish (1996) point out that without guidance, novices can waste large amounts of time (and working memory) creating iterations of maps to figure out the best display parameters.

Online mapping tools on interfaces created for scientists typically pose significant barriers to novices trying to use them. The steps required to create the data visualizations may be confusing to students, and the options overwhelming. Expert terminology used to describe available data parameters can also be very difficult for novices to understand (Edelson, 2005; Edelson, Gordin, & Pea, 1997). The addition of customized scaffolds that explain terminology can provide critically important help, but even in this case students will need to devote significant cognitive resources to selecting appropriate data to include in data visualizations. (See the previous section, ACCESSING DATA, for more on this topic.)

A further challenge for novices arises from the default parameters for data plotting, which are often not appropriate and can lead to confusing data visualizations. For example, the most common default spectral (rainbow) color palette is not a perceptual scale, and application of this palette can obscure important features, highlight unimportant ones, and make it difficult to match a map to its legend (Rheingens & Landreth, 1995; Rogowitz & Treinish, n.d., 1996; Ware 2000, 2008). Because different color palettes are appropriate for different intended uses of a data visualization, providing the option to change the palette is important. However, for novice users to understand how to follow the steps to do this, and what palette to apply when, they will require considerable support and constraining of options beyond what is on the typical scientific data interface (Bergman, Rogowitz, & Treinish, 1995).

Appropriate design choices for geo-referenced data visualizations will vary according to the nature of the data and task. Making design choices for data representations such as maps is not a straightforward process because the design should vary according to the intended use of the visualization and the nature of the data represented (Bergman et al., 1995; Gahegan, 1999; Gahegan & O'Brien, 1997; MacEachren, 2011). The nature of the data and the task dictate the color palette that is most appropriate, the type of interactive features that would be helpful, and the options that should be provided for alternative data displays.

When considering the nature of the task, keep in mind such questions as the following:

- Will the visualization be used for exploration, presentation, and/or communication?
- Will the representation be used to identify patterns in a single data set at a particular time? Changes over time? Divergence from a threshold value?
- Will it be used for making comparisons, such as identifying relationships between different types of data, looking for changes in the same data for two different time periods, or considering the same data from two different regions?
- Will users need to read values and relate them to a legend?

When thinking about the type of data to be represented, consider questions such as the following:

- Are the data 3D? 4D (i.e., time varying)?
- How unusual is the target pattern? (How likely is it that it's on any one map or cross-section?)
- Are the data ordered or categorical?
- What is the spatial frequency of the data?

Because there is no one good design for maps, cross-sections, and 3D displays, it might be tempting to provide a large number of display options in a data interface. However, providing too many options or trying to address too many possible uses can overwhelm a novice user. When making decisions about the design features to include in a particular visualization or the options to provide for students creating data displays, it is important to consider the types of questions that students are most likely to ask and to think about the types of displays that could help them explore these questions (D. Edelson, personal communication, January 11, 2011). The specific guidelines in this section suggest ways to tailor the display to specific data and data uses.

GUIDELINES

USE DESIGN FEATURES THAT ADJUST COGNITIVE LOAD IN ORDER TO MAXIMIZE STUDENTS' ENGAGEMENT IN TASKS RELEVANT TO THE LEARNING GOALS (GUIDELINES 13–17)

As discussed in II. Key Underpinnings: Cognitive Load Theory, to support students' learning from a data visualization, it is important to maximize the degree to which students are engaged in cognitive tasks relevant to the learning goals (germane cognitive load), to simplify the structure of cognitive tasks (i.e., reduce the intrinsic cognitive load), and to minimize engagement in unimportant cognitive tasks (extraneous cognitive load). The following guidelines provide specific suggestions about determining the appropriate amount and type of information to include so that students can work productively with geo-referenced data displays.

Guideline 13: Make visualizations simple and unambiguous, without extraneous information

The British Cartographic Society, in its Five Principles of Map Design (2000), says, "It's not what you put in that makes a great map but what you take out" (slide 10). Cognitive Load Theory (as discussed in II. KEY UNDERPINNINGS: COGNITIVE LOAD THEORY) describes the importance of minimizing the extraneous cognitive load associated with such tasks as reading and finding patterns in a map or cross-section. To promote the allocation of working memory resources to important information, developers should take a number of steps to reduce unnecessary complexity:

- Make sure that data visualizations and other aspects of the interface are uncluttered, that symbols are distinct from one another and their meaning is clear, and that color or graphical elements are used to direct the user's attention to important information (Bunch, 2000; Bunch & Lloyd, 2006; Elvins & Jain, 1998; Nivala, Brewster, & Sarjakoski, 2011; Phillips & Noyes, 1982).
- When possible, have visualizations be task-oriented; remove information that is not relevant to the task (Imhof, 2011).
- Reduce the cognitive load imposed in exploring visualizations of complex data by allowing students to mark and save locations or areas of interest (Plaisant, Carr, & Shneiderman, 1995).
- Reduce the number of data classes in a visualization (thus reducing the number of colors or the number of line symbols in a contour map, for example), which research has indicated can be a straightforward way to reduce demands on working memory (Miller, 1955; Phillips & Noyes, 1982; Slocum, 2008, as cited in Goldsberry & Battersby, 2009).
- Minimize unnecessary clutter in maps by changing the amount of detail displayed as the scale changes; include less detail in smaller-scale maps and place more emphasis on general form (Imhof, 2011).

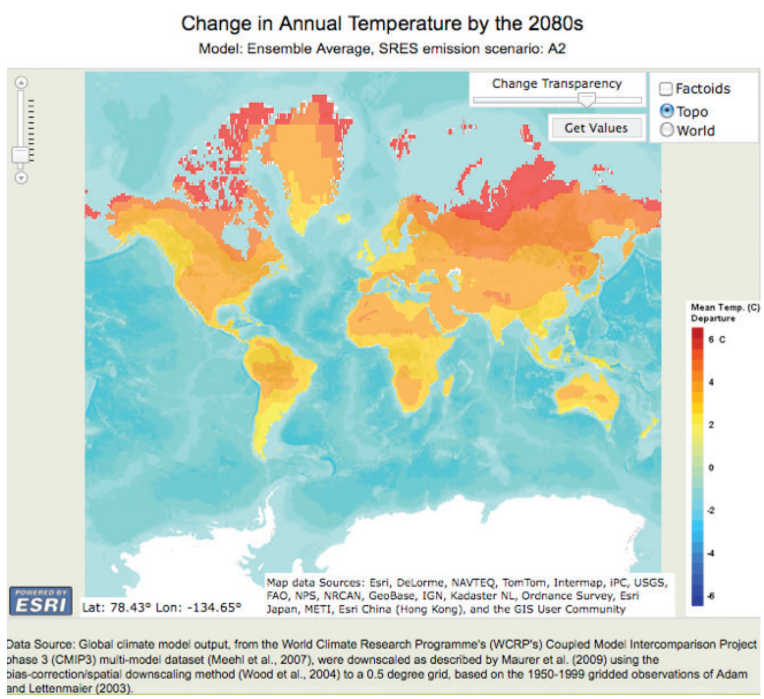


Figure 12. The white area in this map (it is assumed) represents an area where data are not available. However, particularly since it covers Antarctica, students could logically conclude that the white area represents ice. (Other potentially confusing aspects of this map are the blue color of the oceans, which isn't intended to relate to the blue on the color bar, and the way that the map projection greatly exaggerates the size of high-latitude land bodies.) (Source: ClimateWizard, The Nature Conservatory. Retrieved from http://www.climatewizard.org. See the references for this section for additional citations.)

Guideline 14: Include information to minimize confusion and help students establish and maintain orientation

Missing information can make a visualization harder to read and understand, imposing unnecessary demands on working memory. To add clarity to a visualization, consider the following:

- Clearly explain the meaning of colors and symbols.
- Display the units of measurement employed in the visualization.
- Give each visualization a clear title that allows students to quickly identify the type of data they are viewing.
- Include features that orient the student (Edelson, Gordin, & Pea, 1997). For example, physical and political boundaries and labels can allow users to quickly register the location and scale of a map. Sometimes the addition of a simple locus map can be beneficial and help students relate detailed views to more global views. (See also Guideline 32.)

Guideline 15: Identify areas of missing data

Often, areas where data are missing are not adequately explained. For example, white areas designating "no data" on a map may be erroneously confused with ice-covered areas (as shown in Figure 12), and, depending on the color palette used, the areas with no data may be more eye-catching than those with data. Because novices are not likely to be familiar with data collection instrumentation and measurement techniques, they will not anticipate or understand these areas where data are lacking, which may cause confusion and impose unnecessary cognitive load. Therefore, in a visualization, areas where data are missing should be distinct from other areas through the colors or patterns used to designate them, and should be clearly labeled either directly or in an explanatory legend.

Guideline 16: Bridge to the familiar

To reduce the cognitive resources necessary to employ data visualization tools, use icons and features that students are likely to have encountered on other digital interfaces, and make sure they do what they look like they will do (Elvins & Jain, 1998). For example, a magnifying glasses or scale bars containing "+" or "–" symbols , double-tapping, and outward swipes are all commonly used to zoom in on keyboard and tablet computer displays of maps and other information. Inventing new icons or modes of interaction, however clever, may confuse users and distract them from the intended task.

This guideline also applies to other aspects of a geo-referenced data display. Familiar vocabulary and color schemes help to make interactions more intuitive and automatic. For example, the blue ocean color in Figure 12 helps users quickly differentiate land from water, and it is intuitive for red to represent warmer temperatures than green or blue. These aspects of the display make it attractive and user-friendly (even though the color choices are confusing in other ways, as noted in the figure).

Using metaphors can be an effective way to bridge to the familiar. Cartwright (2011) describes nine metaphors for interfaces with geographic information: the Storyteller, the Navigator, the Guide, the Sage, the Data Store, the Fact Book, the Gameplayer, the Theatre, and the Toolbox. The Storyteller metaphor, for example, offers the straightforward option of being "told" information, which doesn't have to be linear. Users can feel like they are in control of the story, but the interface developer can impose the real control by developing and constraining the options. In contrast, the Navigator metaphor can guide users as they navigate through difficult data visualizations by making use of key points or landmarks that users encounter as they explore the visualization. These "landmarks" can be made available sequentially to introduce complex information in steps.

Another example is the desktop metaphor commonly used on computer systems, with screen icons representing features of an office environment (Medyckj-Scott & Blades, 1992). Notation and journaling tools provided to students on data interfaces should reflect the way that students think about and organize their work in school. For example, students could highlight via an icon showing a felt-tip marker, take notes by activating a pencil

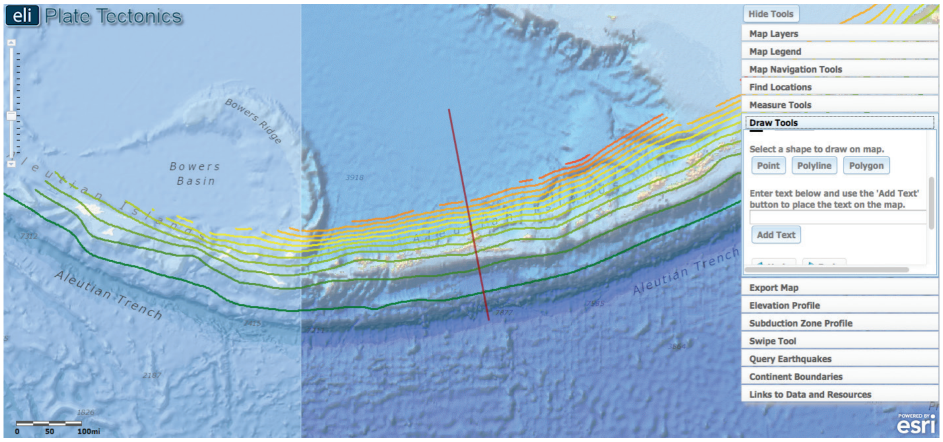


Figure 13. This interactive map provides students with measurement and notation tools, as well as the ability to turn data layers on and off. (Source: Promoting Spatial Thinking with Web-based Geospatial Technologies Project. Retrieved from http://gisweb.cc.lehigh.edu/tectonics/investigation2/.)

icon, and enter and organize notes in an electronic notebook via a notebook icon.

Guideline 17: Minimize and structure text, and effectively integrate it with visualizations

A certain amount of text and labeling is important on a data display to provide explanation and in some cases to highlight important features. However, too much text can distract from the data visualization itself and increase the cognitive resources needed to process the information. For example, if students are required to remember text from a paragraph describing characteristics or features while studying a map, this can overload their working memory. To address this, text should be minimized and visually structured with headings, bullets, etc. so that it is easy to scan for relevant information (Johnson, 2010). Ideally, the text should be integrated into the map or other data visualization; at the very least, the map and text should be presented on the same page and in close proximity to each other (Bunch & Lloyd, 2006; Ware, 2008).

PROVIDE WAYS FOR STUDENTS TO CUSTOMIZE THEIR INTERACTION WITH A GEO-REFERENCED DATA VISUALIZATION (GUIDELINES 18–22)

Guideline 18: Use interactive maps connected to instructional information

Interactive maps can provide information as needed while preserving a simple and clean basic data display. Suggestions and help buttons can provide on-demand guidance to students who need it without adding extraneous cognitive load to the interactions of those who don't (Cartwright, 2011). Interactive maps can also provide a rich background to seemingly simple map renderings by, for example, linking pixels of color to numerical measurements (Edelson, Gordin, & Pea, 1997), linking to photographs or webcams showing buoys or sensors collecting measurements (Cartwright, 1997), or linking to different types of data visualizations (such as drawing a line on a map and linking to a profile view, or linking to data tables).

To provide on-demand access to more in-depth information, Cartwright (2011) suggests using the "Sage" as a metaphor to organize these interactions in a familiar way. Links or hypertext can be used to provide easy access to expert knowledge, such as explanations of expert terminology or descriptions of the instruments used to collect the data.

One caution about relying too heavily on Web links: Because of bandwidth issues and the potential for distraction, it may not be realistic to expect that every student in a classroom will have direct access to the Internet.

Guideline 19: Provide measurement and notation tools

Measurement and notation tools are important to help students keep track of their scientific thinking and bridge from qualitative observations to quantitative measurements, facilitating comparisons and the identification of significant features. Notation tools can allow students to save locations and keep track of areas of a map they have already explored, focus on specific regions for more in-depth exploration, and monitor their own scientific thinking (Plaisant et al., 1995; Quintana et al., 2004). The website developed as part of the Promoting Spatial Thinking with Web-based Geospatial Technologies Project provides tools that students can use to calculate area, measure distances, and mark areas on the map (A. Bodzin, personal communication, July 2, 2012) (see Figure 13).

Guideline 20: Allow data layers to be easily turned on and off

In general, data visualizations should be kept simple, but some patterns are most easily seen when multiple data types are displayed together. Visualizations containing complex information, such as maps, can be scaffolded by allowing the student or teacher to explore information in small, successive segments. For example, in the case of maps, this could involve turning data layers on and off (as in GIS mapping systems, or the interactive map shown in Figure 13). In this way, students and teachers can choose to look at the data sets one at a time before looking at them together (Bunch & Lloyd, 2006; Gahegan, 1999).

Guideline 21: Support students' creation of their own data visualizations

It is particularly important to offer support when students create their own data visualizations. Students' options should be constrained so there are fewer incorrect ways for them to manipulate the interface (Elvins & Jain, 1998), and supports should be included to guide students through the process of selecting display options, such as the appropriate color palette (Harrower & Brewer, 2011; Perkins, Dodge, & Kitchin, 2011). Harrower and Brewer's (2011) ColorBrewer tool provides guidance about the color palettes that work best for different types of map applications, and includes options that accommodate color-blind users. (See Guideline s 23-31 for more specific information about the selection of color palettes for maps and other geo-referenced data visualizations.)

Guideline 22: Allow for more sophisticated interactions as students gain knowledge and skill

How do you build a user interface to data that allows students to enter at different ability levels and then grows with them as their interactions become more sophisticated? Supports and scaffolded steps that may be critical for students when they don't have much relevant prior knowledge or experience can be distracting to and needlessly time-consuming for more experienced students, and data representations that are too simplified won't provide an adequate challenge.

Some suggest that interfaces could be designed to automatically adjust to the prior knowledge, experience, or ability level of the user (Elvins & Jain, 1998; Medyckj-Scott & Blades, 1992). For example, the level of support given to a particular user could adjust to different assumed levels of spatial ability by posing a few simple questions regarding gender, handedness, and experience with science (Bunch & Lloyd, 2006). More sophisticated interfaces could "get to know" a user through his or her interactions. To our knowledge, these types of AI models have not been applied to novices using scientific databases, and much research and development is necessary to make sure that they work effectively. A more reliable way of dealing with this in the science classroom is to allow the teacher, who knows the students, to turn supports on and off.

NRC (2006) suggests providing graded versions of GIS software that are age- and/or experience-appropriate. Others suggest that providing alternative modes can be unnecessarily constraining if students don't fit neatly into a particular mode (W. Finzer, personal communication, June 29, 2011). It is clear that more research is needed in this area.

MAKE THE IMPORTANT INFORMATION AND PATTERNS STAND OUT (GUIDELINES 23–31)

As described in II. KEY UNDERPINNINGS: SCHEMATA, while experts have robust schemata related to the phenomena they are investigating and the patterns they're looking for, novices often lack these schemata and will have difficulty recognizing the scientifically significant patterns buried in the "noise" of complex data visualizations. It is critical for designers to apply what is known about how to use color, textures, shading, and graphical interplay to advantage, ensuring that important features stand out to novice users.

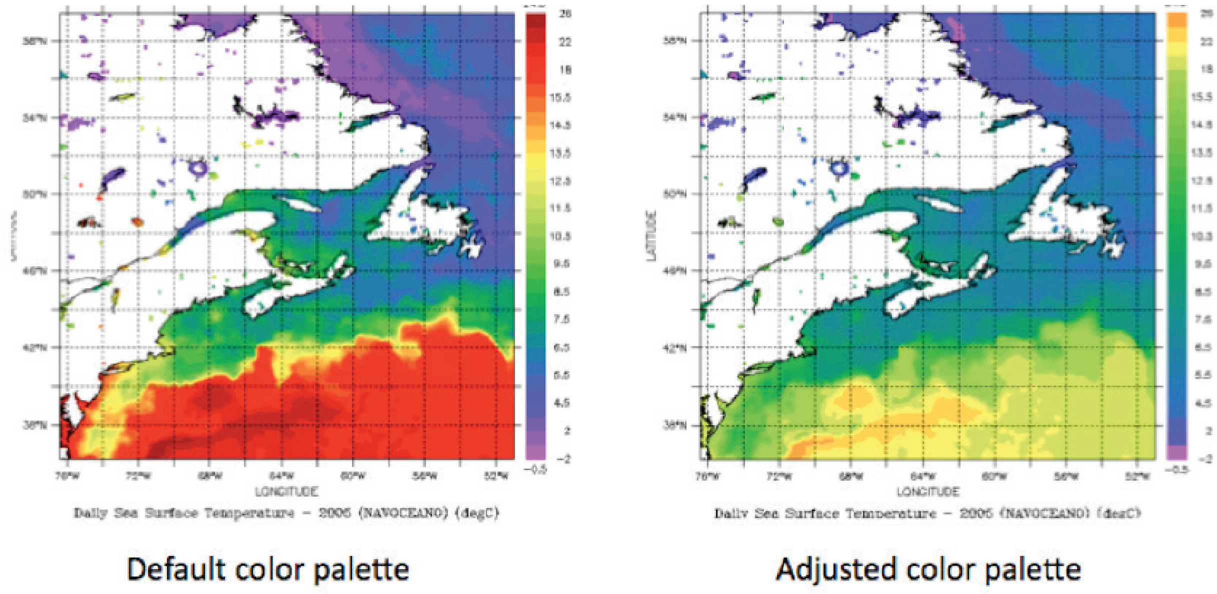


Figure 14. Two different color palettes are used to display the same data. (Source: My NASA Data, National Aeronatics and Space Administration. Retrieved from http://mynasadata.larc.nasa.gov/.)

Guideline 23: Use color to reduce search times for important features

If used correctly, color can be a powerful tool to help students identify significant features, such as important labels, symbols, or data patterns in a map or cross-section. Studies have shown that using color to bring certain features into higher relief results in a better "pop out" effect than differences in other features, such as shape, texture, or orientation (Lloyd, 1997, 2011; Phillips & Noyes, 1980). When a map is to be used to highlight areas over a certain threshold, to differentiate data diverging from a central value (such as above average versus below average rainfall), or to place areas in certain categories that relate to a legend, customized palettes should be used to take advantage of our ability to automatically process certain color differences. (Guidelines 24-30 offer further information about the appropriate palettes to use.)

As described in Cross-Cutting Guideline 2: Draw Attention to Important Features And Patterns, objects or regions that are high in luminance and saturation, that are defined by sharp boundaries, and/or that contrast strongly with other elements of a data display are most likely to catch our eye and direct our attention. However, it's important to make sure that the main elements of a geo-referenced data representation don't perceptually overpower lesser but still important elements, and that the colors applied don't overemphasize certain features so that they mislead the student into seeing more than is actually in the data that are displayed (Imhof, 2011).

Guideline 24: Use color hue to represent categorical (nominal) data

When data are not ordered, but it is important to designate areas or features on a geo-referenced data representation as belonging to different categories, differences in hue work best. For example, you might want to differentiate rivers from roads by using blue for rivers and red for roads, or to designate states as belonging to certain categories that are explained on the legend. If the features are of equal importance, you should keep the level of lightness and saturation equivalent and avoid the perception of ordering (Durstellar, 2008; Harrower & Brewer, 2011; Rogowitz & Treinish, 1996, n.d.).

To make sure that the regions associated with different categories can be easily distinguished and related to a legend, limit the number of categories. As a general rule, cartographers use no more than seven color classes on a map. However, if there are very regular patterns, more than seven color patterns can be used (Harrower & Brewer, 2011).

Guideline 25: For ordered spatial data, use color palettes that vary primarily in luminance and saturation, or use a sequence of hues that are perceptually ordered

Ordered spatial data can range from ordinal data, where the order matters but the difference in value does not, to interval and ratio data, where the difference between two sequential values is meaningful. In all of these cases, it is important for the color sequence to be perceived as ordered. The commonly used rainbow (spectral) palette is not a perceptual scale, meaning, we do not automatically perceive red (at one end of the scale) as greater than purple (at the other end of the scale) or vice versa (Rogowitz & Treinish, 1996, n.d; Ware, 2000, 2008). For interval and ratio data, it is problematic that equal steps in the spectral palette scale are not perceived as equal, and so certain color transitions will stand out as greater and thus more important than others. For example, see Figure 14, which shows the same data displayed using two different color palettes.

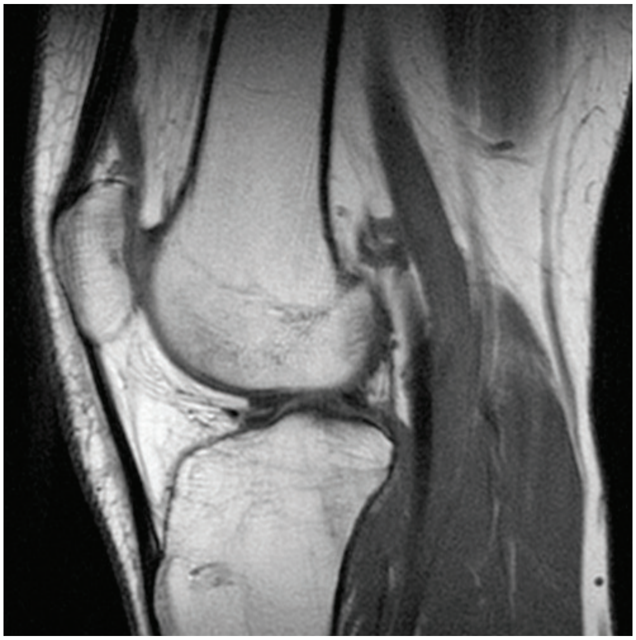


Figure 15. Grayscale is often used in medical imaging, as in this magnetic resonance image of a knee. (Source: Wikipedia, File:MR Knee.jpg. Retrieved from http://en.wikipedia.org/wiki/File:MR_Knee.jpg.)

In the image on the left, which uses the default rainbow palette, the yellow-red region stands out sharply from the blues and greens, and the yellow band seems to be an important boundary. The image on the right, generated using an alternative palette offered on the website, makes the data look quite different, and more accurately portrays the actual changes in temperature. However, notice that the map on the right is still misleading in certain ways. The purple and orange areas, the warmest and coolest areas of the temperature map, stand out as being more significant than the other areas. Unless it is the intended use of the map to highlight areas above and below certain threshold values, this is distracting information.

In ordered color schemes, it is the cartographic convention that lighter colors represent low data values and darker colors represent high values (Harrower & Brewer, 2011). Variations in luminance are particularly good for showing fine detail in ordered data; this is why grayscale is often used in medical imaging, as shown in Figure 15 (Bergman et al., 1995; Durstellar, 2008; Rogowitz & Treinish, 1996, n.d).

For ordered data with lower spatial frequency, variations in saturation (pastel to vivid) work well using only one or a few hues. Hue sequences that are perceptually ordered, combined with changes in luminance, can effectively convey increases or decreases in value, particularly if these color changes relate to our experiences with natural phenomena in the environment. For example, the heated-object scale goes from black through red, orange, and yellow to white, with luminance increasing monotonically. This palette is consistent with human experience, and it conveys more distinguishable display values and contrast between levels than a gray or other single-hued scale (Rheingens & Landreth, 1995).

Note that data visualizations with perceptually continuous color palettes will look more natural because they will convey shape and detail. However, these palettes make it more difficult to match color regions with legends, which is something to consider if it is important for users to derive data values from the map or cross-section. (This may be remedied by allowing the user to see pixel values when clicking on or mousing over different portions of the data display.)

Guideline 26: To show divergence in two directions from an average or threshold value, use two very different hues to represent the extremes, and decrease saturation to a neutral color at the central threshold value

Appropriately designed color palettes can allow the viewer to quickly distinguish data above and below threshold values (Durstellar, 2006; Harrower & Brewer, 2011). Figure 16, for example, shows a palette that might be used to show temperatures above and below an average value, which is represented by "0." The white part of the color bar corresponds to "0" and increases in saturation toward red and blue, which represent increasing divergence above and below the average.

Figure 17 shows regions of Europe that are projected to have increases versus decreases in mean precipitation between now and the end of the century. The more saturated areas draw the viewer's eye to regions where the greatest changes are expected. The saturation changes in a limited number of discrete steps, facilitating the reader's ability to read data values.

Guideline 27: Highlight important thresholds with distinct color changes

When the goal is to highlight data above certain thresholds, customize color palettes so that the data thresholds correspond to perceptually distinct color changes. Use changes in color hue and increases in luminance and saturation to create greater contrast and make the highlighted regions stand out (Bergman et al., 1995; Durstellar, 2008; Ware, 2000, 2008).

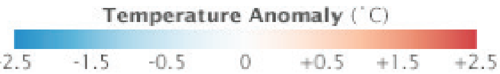


Figure 16. The color bar above shows an example of an appropriately designed color palette for divergent data. (Source: R. Simmon, 2010. NASA Earth Observatory Image of the Day: 2009 Ends Warmest Decade on Record. Retrieved from http://en.wikipedia.org/wiki/File:GISS_temperature_palette.png.)

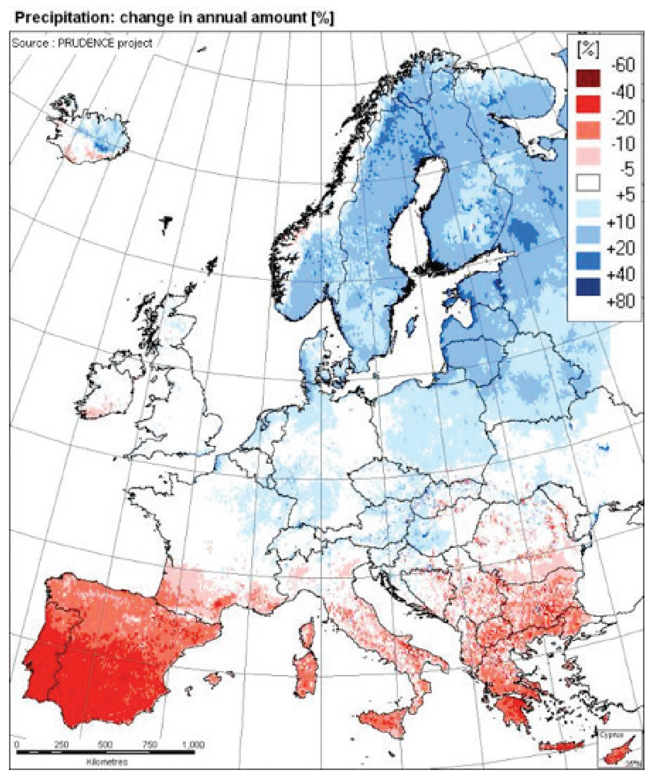


Figure 17. This map displays the change in mean annual precipitation by the end of the 21st century, showing effective use of a divergent color palette. However, a few aspects of this map could be improved: The negative values are near the top of the color bar and the positive values are near the bottom, which is non-intuitive, and the land and ocean are both white and therefore are not clearly differentiated. (Source: The PESETA Project, 2007. Retrieved from http://peseta.jrc.ec.europa.eu/docs/ClimateModel.html.)

Guideline 28: Provide alternative color palettes so that students can customize data representations to the data and task, with clear steps and guidance about how to do so effectively

Since the appropriate color palette to apply to a map or cross-section varies depending on its intended use, it is important to provide alternative palettes for students to use when creating data visualizations. Most data interfaces designed for expert users provide a range of palettes that can be applied. However, the steps to change the palette are not readily apparent to novices, and the alternatives can be confusing. Although students may have fun applying palettes with names such as "beach" and "peppermint" to their maps, they will also likely waste a great deal of time creating attractive or fun maps that don't serve their intended scientific purpose. Alternatively, students may become frustrated and overwhelmed and simply give up. To support students in the creation of useful visualizations, the interface should (1) make it clear how to change the palette and (2) help them identify and apply alternative palettes that are more appropriate to their intended use of the data. This can be achieved through a combination of clearly labeling the palettes, limiting palette options to those most likely to be useful to students, and including guidance on the interface that helps students match a particular palette to their data and intended use (Bergman et al., 1995; Harrower & Brewer, 2011; Healey, St. Amant, & Elhaddad, 1999).

Guideline 29: Design for color-deficient users

Scientists using georeferenced data visualizations rely heavily on color to display patterns. However, approximately 8 percent of men and 0.5 percent of women have color-deficient vision, and it is important for color palettes to be applied to accommodate them (Brewer, MacEachren, Pickle, & Herrmann, 1997; Ware, 2000). Although red-green color-blindness is most common, many other hue pairs can pose difficulties. Color palettes have been developed by researchers to accommodate those with color-deficient vision (see, for example, Brewer & Harrower, 2002). The website Vischeck.com can also be used to check images to determine what they would look like to users with various types of color-deficiencies (Dougherty & Wade, 2008).

In general, red, orange, and green colors may look the same to some students, and it is common for those with color-deficient vision to confuse yellow-green with orange colors (Brewer et al., 1997; Light & Bartlein, 2004). Therefore, use yellow with care (although it will stand out for those with both normal and color-deficient vision) and avoid yellow-green colors altogether. Variations in saturation and luminance will be perceived by those who have color-deficient vision, so they should be used when perceptual ordering is important (Light & Bartlein, 2004). In some cases, varying both hue and saturation or luminance can accommodate those with normal vision as well as those with color-deficient vision.

Guideline 30: Employ design features that minimize confusion caused by color illusions

As discussed in II. KEY UNDERPINNINGS: VISUAL PERCEPTION AND PROCESSING, color can create certain illusions. Perceived hue, luminance, or saturation can be influenced by the color of surrounding objects. The color hue of an object or region can affect its perceived size and distance from the viewer: Red or purple objects and areas appear larger than those that are blue or green; red looks closer than green; and more saturated colors appear closer than less saturated ones. These illusions can make it difficult to match legends with corresponding areas on the data representation and can affect the interpretation of features in important ways (Gahegan, 1999; Harrower & Brewer, 2011; Rheingens & Landreth, 1995; Ware, 2000).

To avoid the pitfalls of these color illusions, pay attention to the outline and background colors used around visualizations and their legends, and limit the hues used in legends to those that can also be clearly distinguished both in the legend and data representation (Harrower & Brewer, 2011; Midtbe, 2001).

Guideline 31: Employ redundancy

Display parameters, such as hue, saturation, lightness, shape, and texture, can often be combined to portray data more effectively. Using two or more stylistic devices at the same time can create greater visual differences between areas, symbols, or labels (Rheingens & Landreth, 1995), as shown in Figure 18.

Redundancy can also be employed to make complex features more readily perceived. For example, Figure 19 uses both hue and boundary lines to highlight water depths less than 750 m (where sound propagates in the ocean) and to distinguish land from ocean. The visualization also employs luminance (light coming from the northwest) to reveal subtle bathymetric features.

The use of multiple display parameters can also help to overcome visual deficiencies among users, such as color-blindness or a lack of vision acuity (see Guideline 29).

A note of caution: Too much redundancy can add extraneous cognitive load to visualizations, and the interplay between graphic elements may also lead to confusion. For example, similar types of patterns laid on top of each other may make both patterns hard to read (Imhof, 2011).

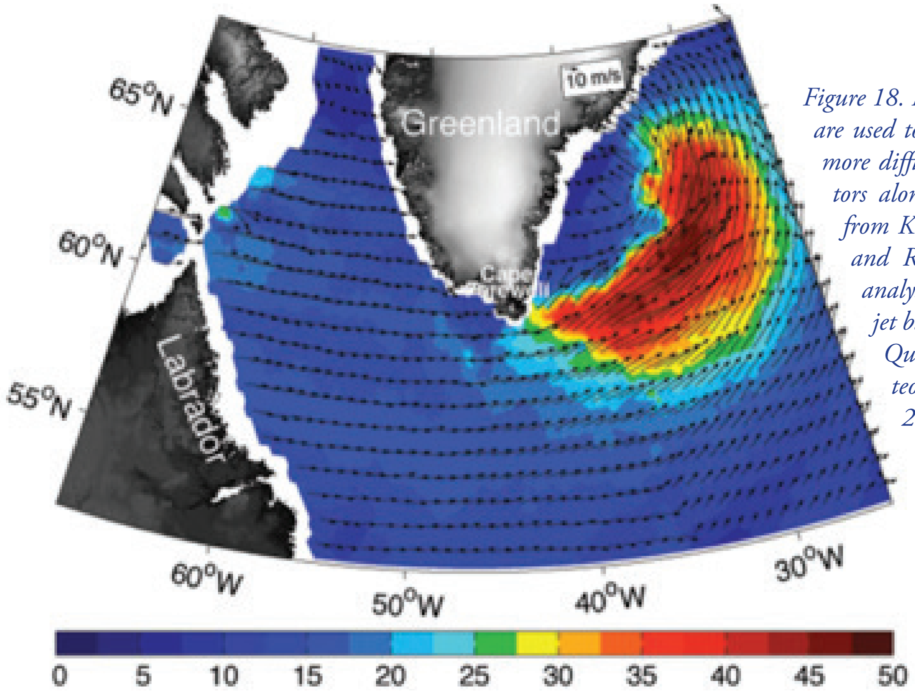


Figure 18. Here, both color and vector length are used to convey magnitude; it would be more difficult to detect patterns using vectors alone. (Source: Modified Figure 2 from K. Våge, T. Spengler, H.C. Davies and R.S. Pickart, 2009. Multi-event analysis of the westerly Greenland tip jet based upon 45 winters in ERA-40. Quarterly Journal of the Royal Meteorological Society, 135, pp. 1999-2011.)

USE TOOLS TO HELP STUDENTS VISUALIZE DIFFERENT PERSPECTIVES AND SCALES (GUIDELINES 32 AND 33)

As mentioned under the CONSIDERATIONS for this section, students will vary significantly in spatial ability, and this will affect the meaning they are able to make of geo-referenced data visualizations. A variety of design features can potentially be employed to support students as they explore data from different perspectives and at different scales.

Guideline 32: Help students relate different map scales by using zoom tools appropriately and by including inset locus maps

Interactive maps that change scale using zoom tools have come into widespread use, so many students will be familiar with them. These tools allow users to move between global views (which show the entire information space) and detailed views. At the time of this writing, common practices include allowing users to draw a box around the region they would like to examine in more detail, or representing the tool with a magnifying glass or with a "+" or "–" symbol. However, although the methods for changing scales may be familiar, students can become disoriented as they move to more detailed views (Plaisant et al., 1995). Using simple, familiar tools and a fixed set of zooming increments, and ensuring that the scale increments are not too great, can be helpful for novice users (Nivala et al., 2011; Plaisant et al., 1995).

To help users coordinate the global and detailed views, Bunch (2000) suggests including a small ancillary or inset window that displays the entire map (and presumably the location of the more detailed view). Plaisant et al. (1995) suggest that scrolling or changing the boundaries of the detailed view should update the global view; they also recommend that the magnification between an overview and a detailed view should be less than 20 if the overview is to be used for navigation. If magnifications greater than this are necessary, intermediate views can be used.

The amount of information displayed should increase as a student zooms to a more detailed view. Google Earth™, for example, does this automatically by adding and subtracting data layers (Geller, 2011). If a student will need to do measurements, a scale bar that automatically adjusts can be included, or zooming can be specified by its factor (such as 150 percent) (Plaisant et al., 1995).

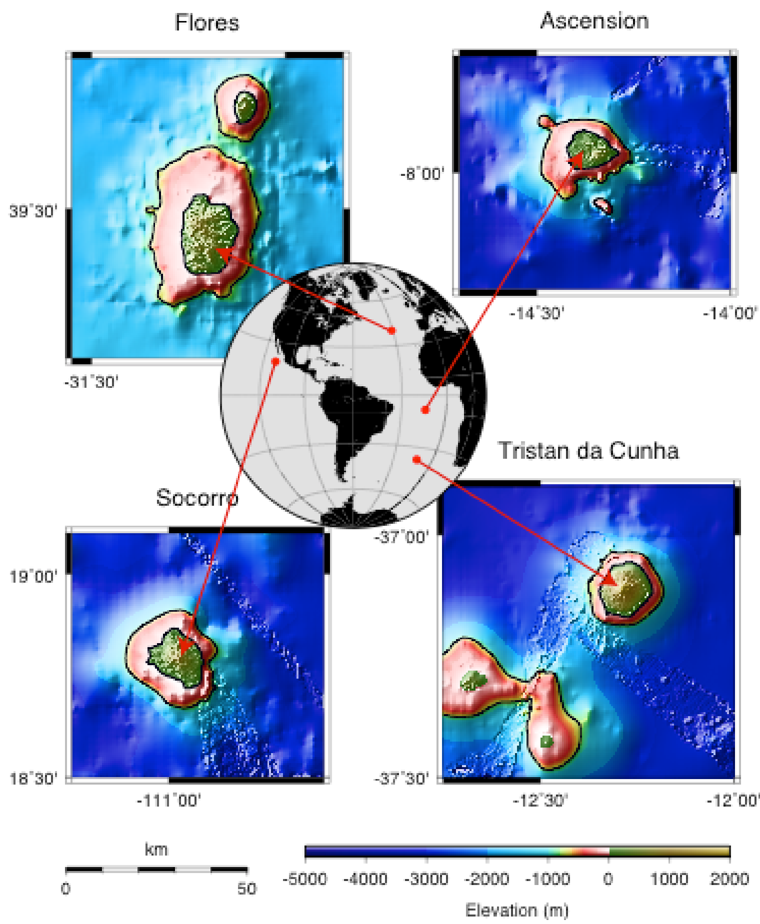


Figure 19. Both hue and boundary lines are used to highlight water depths less than 750 m and to distinguish land from ocean; luminance variations and shading are used to reveal subtle bathymetric features. (Source: Kori Newman plots, BBN Technologies. The elevation data were extracted from the IEDA/MGDS Global Multi-Resolution Topography grid.)

Guideline 33: Use interactive features, 2D and 3D displays, shading, and other visual effects to help students visualize spatial data

Computer interfaces can bring new and promising tools to the table to support students' work with complex, three-dimensional data (Ishikawa & Kastens, 2005). There is much still to learn about which tools will work best for students visualizing complex data. The following are a few suggestions:

- While mental transformation (understanding relationships between views of the same data) is challenging and cognitively demanding (Tory, 2003), allowing users to rotate objects to see surfaces that would otherwise be hidden may ease their cognitive load, particularly for those with low spatial ability (Cook, 2006; Velez, Silver, & Tremaine, 2005).
- When deciding which view to show in static images, choose a viewpoint that contains the most information (Velez et al., 2005).
- There is some evidence that showing both 2D and 3D data visualizations simultaneously is beneficial (Cartwright, 2011).
- Displays that combine 2D and 3D may allow students to determine the 3D position of features more precisely than displays that are either 2D or 3D alone (Tory, 2003; Velez et al., 2005).
- Depth-cueing through variations in color saturation and the use of perspective drawing techniques can aid in 3D visualization (Rheingens & Landreth, 1995).
- Piburn et al. (2002) found that interactive animations, which allowed students to "slice" 3D block diagrams in different ways and to vary their transparency, were effective in helping students visualize the data.
- Virtual reality tools, such as the "fly through" tools provided by Google Earth™, can help students explore and better understand 3D information. However, navigating through environments will not necessarily help students build a cognitive map of the plan view distribution of features (MacEachren, 2004). It has

been suggested that establishing landmarks (distinct features that are visible from a large area) in virtual environments may help users build cognitive maps (Vinson, 1999, cited in Velez et al., 2005).

- Adding shadows can effectively highlight the 3D shapes of surfaces and communicate their relative 3D positioning (Rheingens & Landreth, 1995; Velez et al., 2005).

An interesting research finding is that when adding shadows, the light source should be from the top left (from the northwest on a north-up map view); otherwise, we have trouble perceiving the shape accurately (MacEachren, 2004; Rheingens & Landreth, 1995). This quirk of our visual perception system results from the fact that we are most accustomed to viewing visual scenes with illumination cast from above.

Figure 20 employs a number of the suggested techniques to convey the three-dimensional movements of a satellite-tagged elephant seal.

ENGAGE THE END USER (GUIDELINES 34 AND 35)

Even if great care is taken to adjust cognitive load appropriately, to make the important information stand out, and to provide interactive supports and options for diverse student users, working with authentic scientific data is intrinsically difficult. Novice students in a classroom won't have the same motivation to engage with complex data and visualizations as expert scientists do. Students will bring a variety of attitudes toward science classes and class work. Many (or most!) won't have had experience with the satisfying aspects of working with data and discovering patterns for themselves, and so will be less motivated to push through the "hard parts" of creative problem-solving. While curriculum and instruction offer the most important tools to address the motivation issue, they are beyond the scope of this study. However, the interfaces and tools themselves also have a role to play.

Guideline 34: Provide data and visualization tools that allow students to explore questions they care about

Clearly, an education interface to a scientific database can't and shouldn't provide all possible types of data and data visualization tools—but those provided should support the questions that students are likely to ask and to care about. Thinking through what these questions may be when designing the student interface will help with

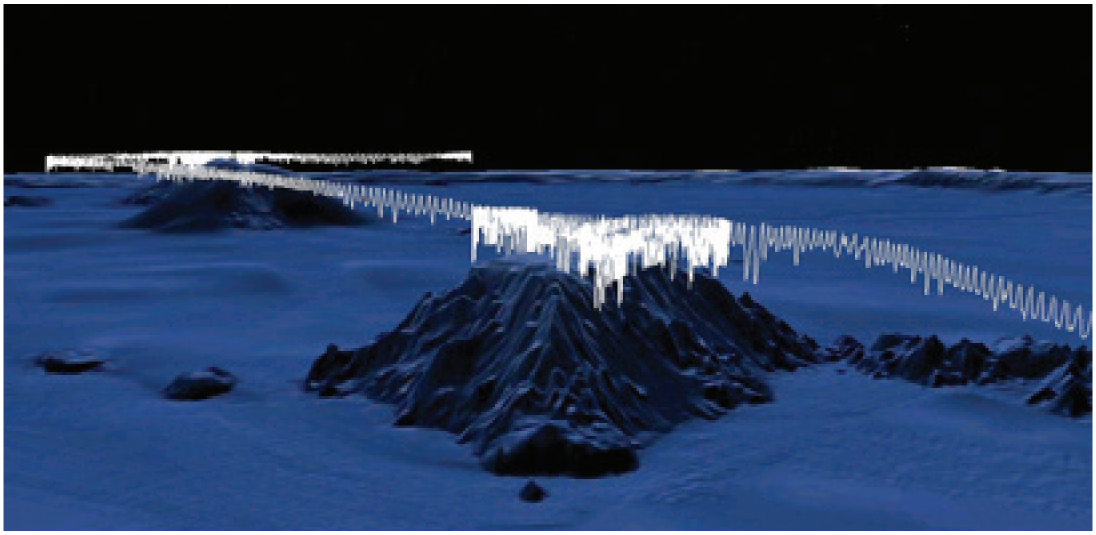


Figure 20. This 3D data visualization shows the track of a satellite-tagged elephant seal superimposed on sea floor bathymetry. The visualization effectively uses luminance differences and lighting from the top left to show detail and convey the 3D shape of the sea floor, contrast to highlight the animal track, and perspective drawing techniques to convey relative distance within the scene. Notice that two data layers are presented, which is key to understanding the pattern of the animal's movement. (Source: Patrick Robinson, UC Santa Cruz, see inside cover for more information)

the decision-making process and ensure that what is provided will support engaging curriculum and instruction (D. Edelson, personal communication, January 11, 2011; Edelson et al., 1997).

Guideline 35: Make the data display aesthetically pleasing to engage the user

> *Engage the emotion to engage the understanding. —British Cartographic Society, The Five Principles of Map Design (2000)*

The overall appearance of a website or data visualization can convey an impression that will affect the users' engagement with the interface—Brewer et al. (1997), for example, found that research subjects preferred more colorful maps. There are indications that higher emotional motivation can increase both students' effort and their engagement with germane cognitive load, thereby increasing learning (Ayres & Youssef, 2008; Bunch & Lloyd, 2006; Cartwright, 2011; Moreno & Mayer, 2000). According to Doering and Veletsianos (2007):

> *Access to geospatial technologies and widely available data sets are spawning a generation of students who "love to explore" using computer software and many times are learning without making a conscious decision that it is time to learn. (p. 223)*

DESIGN FOR ERRORS (GUIDELINES 36 AND 37)

Guideline 36: Help students move beyond errors

Students who are inexperienced with using a scientific database will make errors. Educational interfaces should be designed to minimize possible errors and to make it easier to recover when errors occur (Elvins & Jain, 1998). For example, interfaces could provide clearly worded error messages that tell students what they did wrong and how to correct it, help links, back buttons, easy ways to save images, and reminders to do so (Nivala et al., 2011; Shneiderman, 1998). Students should be able to easily undo any actions they take, which will reduce their anxiety and give them the confidence to explore (Shneiderman, 1998).

Guideline 37: Help students recognize data problems

Plots generated by an online database will reflect instrument and processing malfunctions. Students generating such plots may not recognize data problems that would cause a scientist to throw out a plot; they may instead interpret measurement errors as real features. For example, students attempting to understand average air temperatures during the summer along the coast of Maine may not recognize the significance of smooth contours breaking up into boxes in Figure 21—a consequence of data resolution problems along the boundary of where data are available.

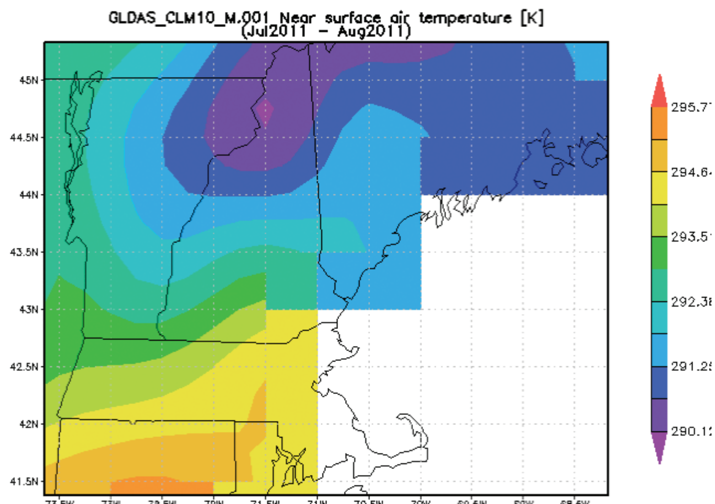


Figure 21. This map, generated by a teacher at a school in coastal Maine, exhibits problems associated with data resolution along the coast. (Source: Generated via Goddard Earth Sciences Data and Information Services Information Center, http://disc.sci.gsfc.nasa.gov/giovanni/overview/index.html.)

Nonetheless, these problems with data quality do present important "teachable" moments to students, because they remind students to think about the data source, which is remote from the students, and the data filtering and conditioning that occurs between the sensors and the interface provided to student users (Gould, Sunbury, & Krumhansl, 2012).

PROVIDE GEO-REFERENCED DATA VISUALIZATIONS AND TOOLS THAT SUPPORT THE TEACHING OF SCIENTIFIC PRACTICES (GUIDELINES 38 AND 39)

The goal of engaging students in work with authentic scientific databases is to give them opportunities to develop their understanding of scientific practices. Thus, students should be provided with tools that will allow them to perform data analyses that are typically performed by scientists and to investigate questions of scientific importance.

Guideline 38: Pre-determine the types of visualizations that students can create based on what scientists typically use to explore a given issue

As described in other guidelines, cognitive load can be beneficially adjusted by limiting the choices that are provided to students. When deciding how to limit the ways that students may view the data, choose options that scientists themselves typically use to view certain data sets (Gordin, Polman, & Pea, 1994). For example, with respect to maps, wind can be displayed as vectors, and temperature can be displayed as a contour map. By using the same types of displays typically used by scientists, as pointed out by Gordin et al. (1994), "these inscriptions can serve as vehicles that allow students to legitimately participate in the community of scientists which also uses these representations" (p. 216).

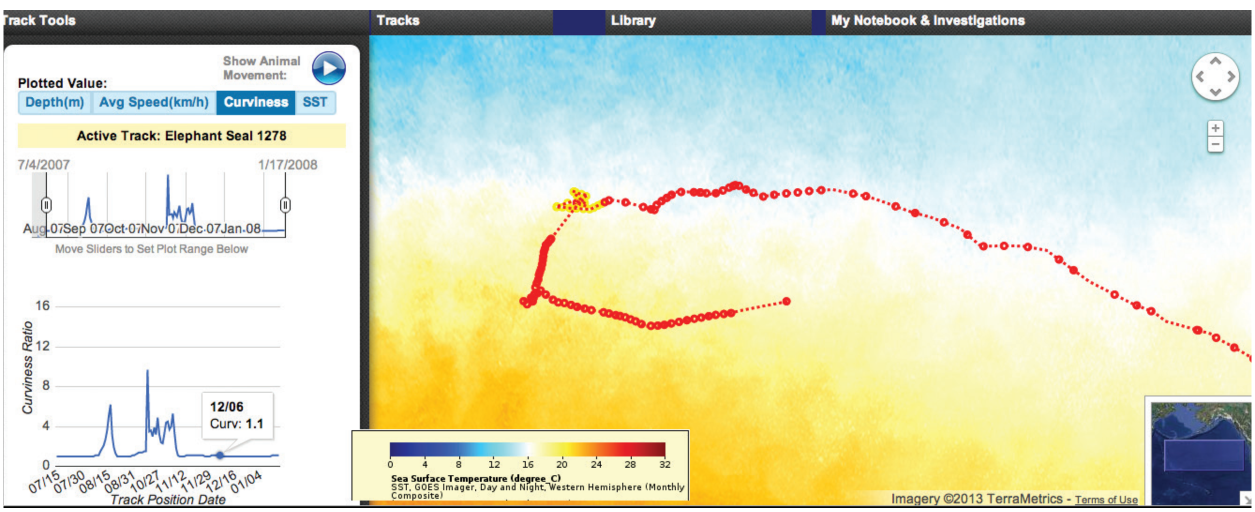


Figure 22. A simple tool that allows students to quantitatively measure and compare the curviness of different lengths of track is being created as part of the NSF-funded Analyzing Ocean Tracks Project. The red line in the map shows the path of a satellite-tagged elephant seal in the Pacific Ocean. The tortuosity (curviness) of marine animal tracks provides clues to the animals' behavior; track segments that are more curvy may indicate feeding, while straighter tracks are indicative of migration. A tool will allow students to measure and average the angles in the track shape to facilitate comparison of track segments. Notice that the map also shows sea surface temperature—a data layer that students can turn on and off to examine the relationship of animal tracks to thermal fronts and ocean currents. (Source: Analyzing Ocean Tracks Project. Retrieved from http://oceantracks.org).

Guideline 39: Provide tools and settings that allow for easy comparison of data visualizations and quantitative analysis of features

To answer many scientific questions, it is necessary to go beyond qualitative observations of a single data map or cross-section. For example, one might need to make comparisons between maps of the same data at different time periods, or maps of different but potentially related data sets at the same time period. Educational interfaces to scientific databases should facilitate these types of comparisons. Layering data on the same map can make comparisons easier; programs such as GIS and Google Earth™ allow the user to choose layers and turn them on and off (Geller, 2011). The Promoting Spatial Thinking with Web-based Geospatial Technologies Project (Bodzin, n.d.) also provides such tools (see Figure 13).

Gordin et al. (1994) suggest that having the scale of a map automatically set as a default can simplify students' interaction with an online database. However, as Gordin et al. also recognized, these default settings don't work for every application. To compare the same type of data—for example, near-surface air temperature—in the same area for two different time periods, the scales and meaning of colors in contoured data should be kept constant so that the colors refer the same temperature range in the two different data representations. Some scientific data interfaces automatically adjust the color bar for each generated visualization. Although this optimizes the amount of data structure (i.e., the variability within the map or cross-section) that can be displayed, it may confuse students when they try to do comparisons. Perhaps having clearly labeled display settings for comparing geo-referenced data representations (versus showing the data structure in individual maps or cross-sections) would help students think scientifically about the data analyses they intend to do, and avoid the frustration of working with visualizations that aren't displaying data in meaningful or useful ways.

Scientists also perform other operations on mapped data, such as subtracting two data sets or quantitatively analyzing the shape of certain features (Gordin & Pea, 1995). The inclusion of simplified tools that students can use to perform such operations (as described in Figure 22) will help them build sophistication in their analytical practices. As students gain experience, they will encounter situations in which the built-in tools don't help them answer the questions they have. Therefore, an interface should give them an opportunity to invent and name their own measures, and provide them with support in doing so (W. Finzer, personal communication, December 5, 2011).

REFERENCES

Ayres, P., & Youssef, A. (2008). Investigating the Influence of Transitory Information and Motivation During Instructional Animations. *ICLS '08, Proceedings of the 8th International Conference of the International Society of the Learning Sciences, 1,* 68-75.

Bednarz, S. W. (2001). *Thinking spatially: Incorporating geographic information science in pre and post secondary education.* Retrieved from http://www.geography.org.uk/download/EVbednarzthink.doc

Bergman, L. D., Rogowitz, B. E., & Treinish, L. A. (1995). A rule-based tool for assisting colormap selections. In *Proceedings of the IEEE Visualization 1995 Conference* (pp. 118-125). New York, NY: IEEE.

Bodzin, A. M. (n.d.). *Promoting Spatial Thinking with Web-based Geospatial Technology.* Retrieved from http://coe.lehigh.edu/research/promoting-spatial-thinking-web-based-geospatial-technology

Brewer, C. A., & Harrower, M. (2002). *ColorBrewer 2.0: Color Advice for Cartography.* Retrieved from http://colorbrewer2.org/

Brewer, C. A., MacEachren, A. M., Pickle, L. W., & Herrmann, D. (1997). Mapping mortality: Evaluating color schemes for chloropleth maps. *Annals of the Association of American Geographers, 87*(3), 411-438.

British Cartographic Society. (2000). *Cartographic Principles: A Perspective from the British Cartographic Society Design Group* [PowerPoint presentation]. Retrieved from http://www.geo.wvu.edu/~elmes/geog462/cartographicprinciples.pdf

Bunch, R. L. (2000). GIS and the acquisition of spatial information: Differences among adults and young adolescents. *Research in Geographic Education, 2*, 67-97.

Bunch, R. L., & Lloyd, R. E. (2006). The cognitive load of geographic information. *The Professional Geographer, 58*(2), 209-220.

Cartwright, W. (1997). New media and its application to the production of map products. *Computers and Geosciences, 23*, 447-456.

Cartwright, W. (2011). Extending the map metaphor using Web delivered multimedia. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader* (chapter 2.11). Chichester, UK: Wiley-Blackwell.

Chabris, C. F., Jerde, T. E., Woolley, A. W., Gerbasi, M. D. Schuldt, J. P., Bennett, S. L., Hackman, J. R., & Kosslyn, S. M. (2006). *Spatial and object visualization cognitive styles: Validation studies in 3800 individuals.* Retrieved from http://www.wjh.harvard.edu/~cfc/Chabris2006d.pdf

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education, 90*, 1073-1091.

Doering, A., & Veletsianos, G. (2007). An investigation of the use of real-time, authentic geospatial data in the K-12 classroom. *Journal of Geography, 106*(6), 217-225.

Dougherty, R. F., & Wade, A. R. (2008). *Vischeck.com.* Retrieved from http://www.vischeck.com

Durstellar, J. C. (2006). *Colour guidelines.* Retrieved from http://www.infovis.net/printMag.php?num=184&lang=2

Durstellar, J. C. (2008). *The spectral (or rainbow) palette.* Retrieved from http://www.infovis.net/printMag.php?lang=2&num=192

Edelson, D. (2005, February 16-19). E*ngineering Pedagogical Reform: A Case Study of Technology Supported Inquiry. NSF Inquiry Conference on Developing a Consensus Research Agenda,* Rutgers University, New Brunswick, NJ.

Edelson, D. C., Gordin, D., & Pea, R. D. (1997, March 20-24). *Creating Science Learning Tools from Experts' Investigation Tools: A Design Framework.* Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Oak Brook, IL.

Elvins, T. T., & Jain, R. (1998). Engineering a human factor-based geographic user interface. *IEEE Computer Graphics and Applications, 18*(3), 66-77.

Gahegan, M. (1999). Four barriers to the development of effective exploratory visualization tools for the geosciences. *Geographical Information Science, 13*(4), 289-309.

Gahegan, M. N., & O'Brien, D. L. (1997). A strategy and architecture for the visualization of complex datasets. *International Journal of Pattern Recognition and Artificial Intelligence, 11*, 239-261.

Geller, T. (2011). Imaging the world: The state of online mapping. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader* (chapter 2.12). Chichester, UK: Wiley-Blackwell.

Goldsberry, K., & Battersby, S. (2009). Issues of change detection in animated chloropleth maps. *Cartographica, 44*(3), 201-215.

Gordin, D. N., & Pea, R. D. (1995). Prospects for scientific visualization as an educational technology. *The Journal of the Learning Sciences, 4*(3), 249-279.

Gordin, D. N., Polman, J. L., & Pea, R. D. (1994). The Climate Visualizer: Sense-Making Through Scientific Visualization. *Journal of Science Education and Technology, 3*(4), 203-226.

Gould, R. R., Sunbury, S., & Krumhansl, R. (2012, May). Using online telescopes to explore exoplanets from the physics classroom. *American Journal of Physics, 80*(5), 445-451.

Harrower, M., & Brewer, C. A. (2011). ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader* (chapter 3.8). Chichester, UK: Wiley-Blackwell.

Healey, C. G., St. Amant, R., & Elhaddad, S. (1999, October 11-13). *ViA: A perceptual visualization assistant.* Conference paper presented at the 28th Workshop on Advanced Imagery Pattern Recognition, Washington, DC.

Imhof, E. (2011). Interplay of Elements, from Cartographic Relief Presentation. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader* (chapter 3.2). Chichester, UK: Wiley-Blackwell.

Ishikawa, T., & Kastens, K. (2005). Why some students have trouble with maps and other spatial representations. *Journal of Geoscience Education, 53*(2), 184-197.

Johnson, J. (2010). *Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules.* Burlington, MA: Morgan Kaufmann.

Kali, Y., & Orion, N. (1996). Spatial abilities of high-school students in the perception of geologic structures. *Journal of Research in Science Teaching, 33*, 369 391.

Kastens, K. (2010, March). Commentary: Object and Spatial Visualization in Geosciences. *Journal of Geoscience Education, 58*(2), 52-57.

Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory and Cognition, 33,* 710 726.

Light, A., & Bartlein, P. J. (2004). The end of the rainbow? Color schemes for improved data graphics. *Eos, 85*(5), 385-391.

Lloyd, R. (1997). Visual search processes used in map reading. *Cartographica, 34*(1), 11-31.

Lloyd, R. (2011). Understanding and learning maps. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader* (chapter 4.9). Chichester, UK: Wiley-Blackwell.

MacEachren, A. M. (2004). *How Maps Work: Representation, Visualization, and Design.* New York, NY: The Guilford Press.

MacEachren, A. M. (2011). The Roles of Maps, from Some Truth with Maps: A Primer on Symbolization and Design. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader* (chapter 3.6). Chichester, UK: Wiley-Blackwell.

Medyckj-Scott, D., & Blades, M. (1992). Human spatial cognition: Its relevance to the design and use of spatial information systems. *Geoforum, 23*(2), 215-226.

Midtbe, T. (2001, August 6-10). Perception of time variables on dynamic maps. In *Proceedings of the 20th International Cartographic Conference*, Beijing, Peoples Republic of China.

Miller, G. A. (1955). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review, 101*(2), 34-352.

Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology, 92*(1), 117-125.

National Research Council. (2006). *Learning to Think Spatially: GIS as a Support System in the K-12 Curriculum.* Washington, DC: The National Academies Press.

Nivala, A. -M., Brewster, S., & Sarjakoski, L. T. (2011). Usability Evaluation of Web Mapping Sites. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader* (chapter 4.11). Chichester, UK: Wiley-Blackwell.

Perkins, C., Dodge, M., & Kitchin, R. (2011). Introductory Essay: Cartographic aesthetics and map design. In M. Dodge, R. Kitchin, & C. Perkins (Eds.), *The Map Reader* (chapter 3.1). Chichester, UK: Wiley-Blackwell.

Phillips, R. J., & Noyes, L. (1980). A comparison of colour and visual texture as codes for use as area symbols on thematic maps. *Ergonomics, 23*(12), 1117-1128.

Phillips, R. J., & Noyes, L. (1982). An investigation of visual clutter in the topographic base of a geological map. *The Cartographic Journal, 19*(2), 122-132.

Piburn, M., Reynolds, S., Leedy, D., McAuliffe, C., Birk, J., & Johnson, J. (2002, April 7-10). *The Hidden Earth: Visualization of Geologic Features and Their Subsurface Geometry.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Plaisant, C., Carr, D., & Shneiderman, B. (1995). Image-browser taxonomy and guidelines for designers. *IEEE Software, 12*(2), 21-32.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D., & Soloway, E. (2004). A scaffolding design framework for software to support scientific inquiry. *The Journal of the Learning Sciences, 13*(3), 337-386.

Rheingens, P., & Landreth, C. (1995). Perceptual principles for effective visualizations. In G. Grinstein & H. Levkowits (Eds.), *Perceptual Issues in Visualization* (pp. 59-74). Berlin, Germany: Springer-Verlag.

Rogowitz, B. E., & Treinish, L. A. (1996, May/June). How Not to Lie with Visualization. *Computers In Physics, 10*(3), 268-273.

Rogowitz, B. E., & Treinish, L. A. (n.d.). *Why Should Engineers and Scientists Be Worried About Color?* Yorktown Heights, NY: IBM Thomas J. Watson Research Center. Retrieved from http://www.research.ibm.com/people/l/lloydt/color/color.HTM

Shneiderman, B, (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (3rd ed.). Reading, MA: Addison-Wesley.

Svenson, S., & Kastens, K. (2011). Student interpretation of a global elevation map: What it is, how it was made, and what it is useful for. In A. D. Feig & A. Stokes (Eds.), *Qualitative Inquiry in Geoscience Education Research: Geological Society of America Special* (pp. 189-211). Boulder, CO: The Geological Society of America.

Tory, M. (2003). Mental registration of 2D and 3D visualization (an empirical study). In *Proceedings of the 14th IEEE Visualization Conference* (pp. 371-378). Seattle, WA: IEEE.

Velez, M. C., Silver, D., & Tremaine, M. (2005, October). Understanding visualization through spatial ability differences. In *Proceedings of the IEEE Visualization 2005 Conference* (pp. 511-518). New York, NY: IEEE.

Ware, C. (2000). *Information Visualization: Perception for Design.* San Francisco, CA: Morgan Kaufmann.

Ware, C. (2008). *Visual Thinking for Design*. Burlington, MA: Morgan Kaufmann.

CITATIONS FOR FIGURE 12:

Adam, J. C., & Lettenmaier, D. P. (2003). Adjustment of global gridded precipitation for systematic bias. *Journal of Geophysical Research, 108*, 1-14.

Maurer, E. P., Adam, J. C., & Wood, A. W. (2009). Climate model based consensus on the hydrologic impacts of climate change to the Rio Lempa basin of Central America. *Hydrology and Earth System Sciences, 13*, 183-194. Retrieved from http://www.hydrol-earth-syst-sci.net/13/183/2009/hess-13-183-2009.html

Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., Stouffer, R. J., & Taylor, K. E. (2007). The WCRP CMIP3 multi-model dataset: A new era in climate change research. *Bulletin of the American Meteorological Society, 88*, 1383-1394.

Wood, A. W., Leung, L. 5 R., Sridhar, V., and & Lettenmaier, D. P. (2004).: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs., *Climatic Change, 62*, 189-216, 2004.

Graphs

> *Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the shortest space.*
>
> *—Edward R. Tufte, The Visual Display of Quantitative Information, 1983 (p. 51)*

CONSIDERATIONS

WHAT ARE WE TALKING ABOUT?

Graphs are one of the most important ways that scientists communicate and interpret relationships among data. Making graphs is a ubiquitous practice among scientists (Latour, 1990). But why are graphs so central to scientific practice?

A graph both transforms and reduces large amounts of data into a visual representation, which can then communicate patterns in data that typically cannot be directly visualized, due to the nature of the phenomena being studied or the number of data points that are being sampled. Graphs use spatial and visual features, such as length, angle, area, and color, to represent quantitative and categorical data and to show the relationships among data. The references for the data are usually, but not always, represented on axes. Graphs are used to analyze, explain, and predict phenomena. They are most commonly used to show a change in a variable against a change in time, but they can show other relationships as well.

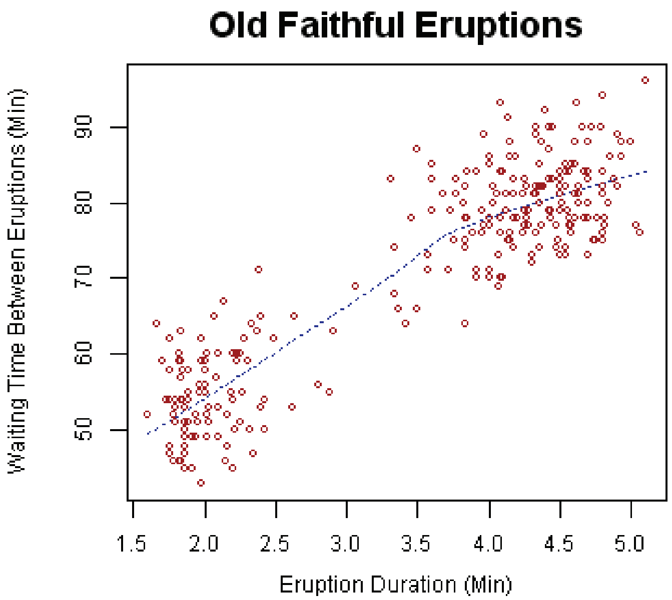


Figure 23. This scatter plot compares the waiting time between eruptions of the geyser Old Faithful to the eruption duration. This graph makes two patterns clear: (1) Eruptions seem to fall into two groups—longer duration and shorter duration, with few in the middle and (2) there is more wait time between the longer-duration eruptions. (Source: Wikipedia, File:Oldfaithful3.png. Retrieved from http://en.wikipedia.org/wiki/File:Oldfaithful3.png.)

Graphs can also show statistical measures, which relate to the collection, organization, and interpretation of numerical data, and often describe a specific kind of summary of the individual data points in a data set. Statistical measures that are appropriate for high school students to consider are trend lines (also called "regression lines," "lines of best fit," or "least squares lines") and measures of center, such as average.

An educational data interface might feature many types of graphs including scatter plots, line graphs, grouped line graphs, bar graphs, pie charts, box plots, anomaly charts, and "bubble charts" (a variant of a scatter plot in which the size of the data points—the bubbles—relay quantitative information about a variable). Each type is useful for a specific purpose, for example:

- Relating quantitative data sets
- Relating a quantitative data set to a categorical data set
- Showing the proportion of one data set to another data set
- Showing the variability in a data set

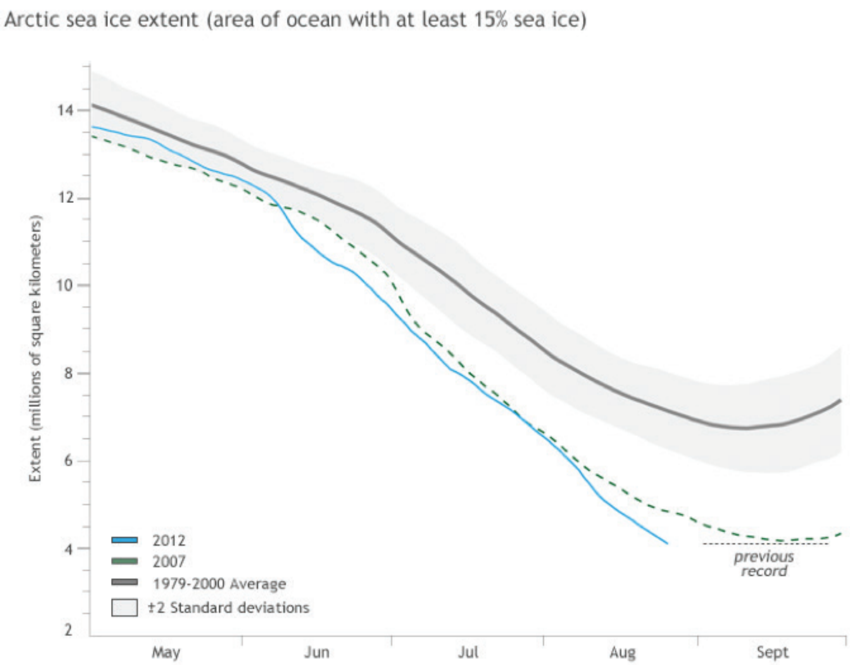


Figure 24. Line graphs of arctic sea ice against month for 2007 and 2012. An average is also included, for purposes of comparison. (Source: R. Lindsey, 2012. Arctic Sea Ice Breaks 2007 Record Low. ClimateWatch Magazine. Retrieved from http://www.climate-watch.noaa.gov/article/2012/arctic-sea-ice-breaks-2007-record-low.)

A scatter plot usually shows data points that represent two variables in a coordinate plane, with one variable referenced on the horizontal axis and the other variable referenced on the vertical axis. Scatter plots are particularly useful for showing trends in large data sets, including the real-time and archived observatory data that is newly available to students online. Figure 23 shows a scatter plot of the waiting time between eruptions of the geyser Old Faithful and the duration of eruptions.

A line graph is similar to a scatter plot in that it shows data points that represent two or more variables in a coordinate plane; however, the data points are connected to each other with a line. Figure 24 shows a grouped line graph, which contains multiple line graphs in one representation.

A bar graph or histogram, such as the simple one of tree height in Figure 25, shows data on two axes by using parallel bars or rectangles. This type of graph is often used with categorical data, which groups or "bins" a number of individual data points into a category.

WHAT SHOULD THE USER INTERFACE DESIGNER CONSIDER?

Most students do not understand the nature of data. Because the spatial and visual features in graphs show relationships among data rather than map directly to any physical place or process, a robust understanding of data in general, and of data sets plotted in graphs in particular, are important to support graph comprehension. However, students have difficulty in understanding what a data set represents (Konold, Higgins, Russell, & Khalil, 2004). A data set contains measurements

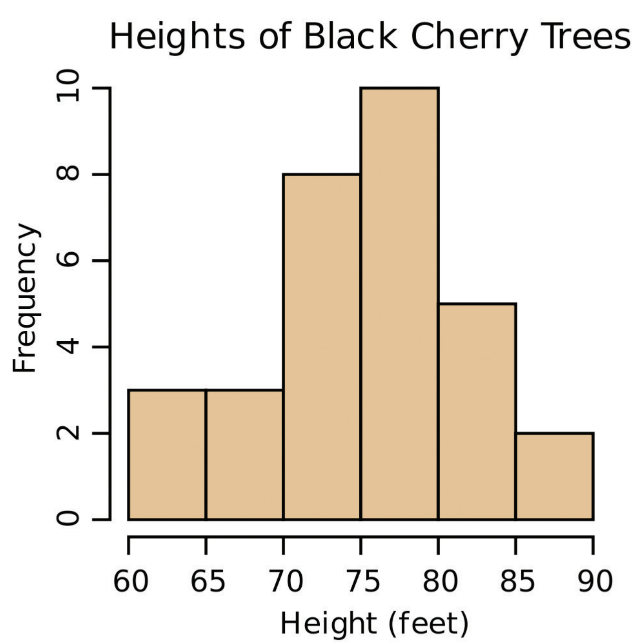


Figure 25. Bar graph of the heights of cherry trees against frequency (Source: Wikipedia, File:Black cherry tree historgram.svg. Retrieved from http://en.wikipedia.org/w/index.php?title=File:Black_cherry_tree_histogram.svg&page)

of objects or events that have been grouped together. The basis of this grouping may be obscure or puzzling to students, as the objects or events may be alike only in one attribute and differ in all other attributes. For example, the amount of rainfall in a desert and in a rain forest can be compared, even though these two places have little in common. A related difficulty that students have is a limited understanding of measures of center, such as mean or average, median, and mode (Hammerman & Rubin, 2004; Konold et al., 2004; Konold & Pollatsek, 2002; Mokros & Russell, 1995). A measure of center is an important concept in data because it provides a means for characterizing all the values in a data set with just one number.

Graph comprehension consists of many kinds of cognitive tasks, and the complexity of the process imposes high intrinsic cognitive load. A number of researchers have studied and described the graph comprehension process (Carpenter & Shah, 1998; Gillian & Lewis, 1994; Ratwani, Trafton, & Boehm-Davis, 2008). This process is complex, iterative, and requires many abilities, including spatial ability. The steps of graph comprehension include the following, though their order may vary:

- The student notes the visual features of a graph and then separates those features into a series of visual chunks of information. A visual chunk could be the entire line in a line graph, a section of a line in a line graph with a particular slope, a bar in a bar graph, two grouped bars in a grouped bar graph, a slice in a pie graph, and so on.
- The student studies the visual chunk to recognize its visual pattern.
- The student integrates the visual pattern of a chunk with the textual information on the graph, such as labels and numbers, and interprets the data relationship.
- The student completes the same cycle of identification, visual pattern recognition, integration, and interpretation for each additional visual chunk of information on the graph.
- The student combines his or her interpretations of all the visual chunks of information in order to determine the meaning of the entire graph.

Eye movement research has shown that when students integrate the visual pattern of a visual chunk with textual information, they reexamine the visual chunks, axis labels, scales, and other labels, such as keys or legends, multiple times. This repetition suggests that students cannot keep all the information in the graph in their working memory, indicating that graph comprehension imposes high intrinsic cognitive load. (For more information on cognitive load, see II. Key Underpinnings: Cognitive Load Theory.)

Accuracy of visual perception varies according to the visual features found in a graph. Students use their spatial and visual abilities to interpret the data relationships of visual chunks in a graph, and their interpretation is based partly on their ability to evaluate particular visual features of the graph—what researchers call a "graphical-perception task." One kind of graphical-perception task is the cognitively automatic extraction of quantitative information. (For more information, see II. Key Underpinnings: Visual Perception and Processing.) Cleveland and McGill (1984) studied this kind of graphical-perception task in depth and determined that five specific visual features—position, shape, size, symbols, and color—and nine graphical-perception tasks—differentiating area, color hue, color saturation, density, length, position along a common scale, position on identical but nonaligned scales, slope, and volume—are used to decode a graph's quantitative information. Students' accuracy in automatically extracting quantitative information from

RANK	VISUAL FEATURE
1	Position along a common scale
2	Position on identical but nonaligned scales
3	Length
4	Angle Slope (with θ not too close to 0, $\pi/2$, or π radians)
5	Area
6	Volume Density Color saturation
7	Color hue

Table 1. Accuracy of extracting quantitative information, ordered from most accurate to least.

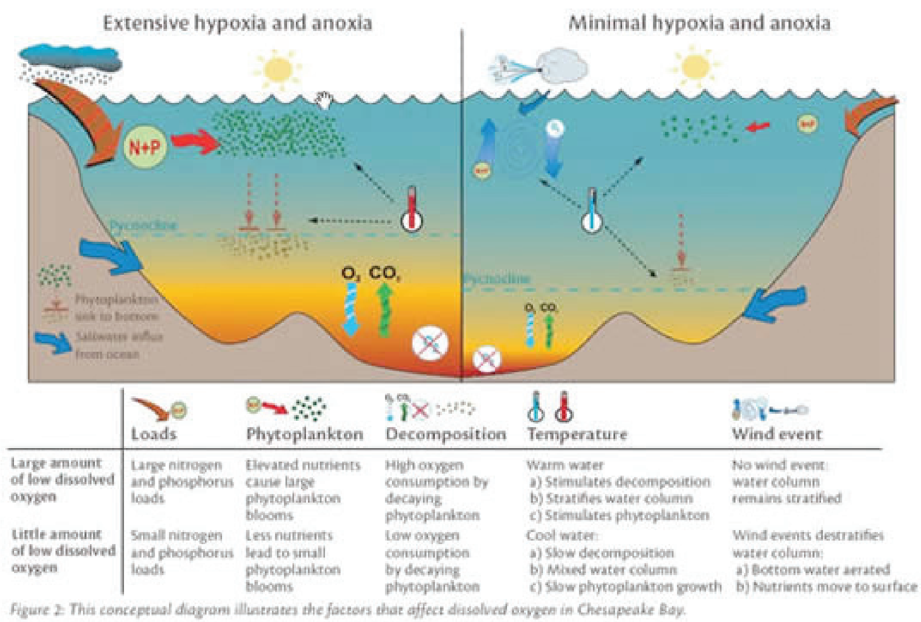


Figure 26. This diagram shows how data connect to an Earth process. (Source: Goddard Earth Sciences and Data Information Center, National Aeronautics and Space Administration. Retrieved from http://disc.sci.gsfc.nasa.gov/education-and-outreach/additional/science-focus/locus/images/DO_diagram.jpg)

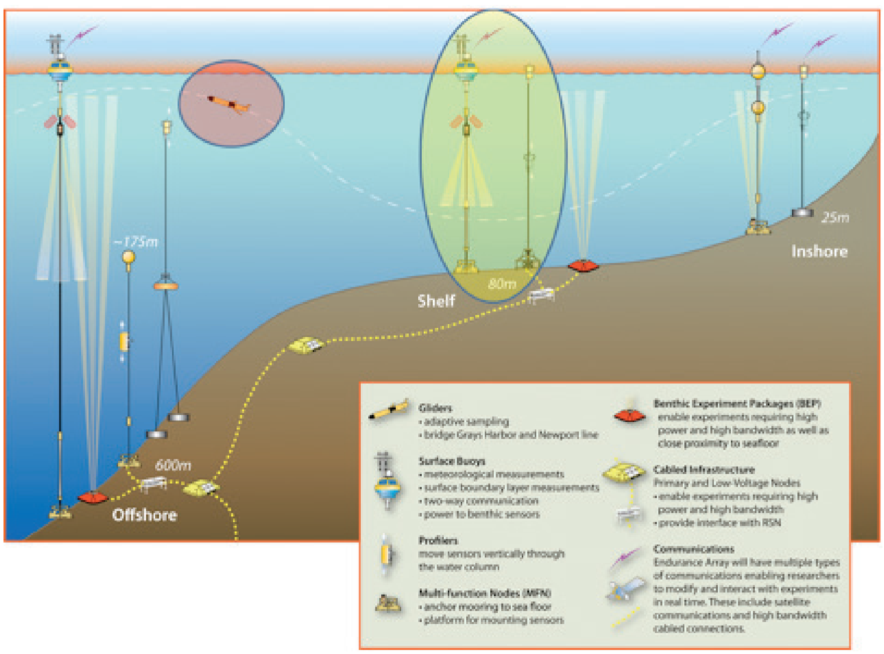


Figure 27. This diagram describes how dissolved oxygen data are collected with ocean sampling instruments, as part of the Ocean Observatories Initiative. (Source: Illustration by David Reinert, copyright Oregan State University)

particular visual features will vary according to the kind of feature that is being decoded. Table 1, adapted from Cleveland and McGill, ranks the accuracy of extracting quantitative information.

The interpretation and creation of graphs by novice students will not be the same as an expert's. In contrast to experts, most novices lack robust schemata that can help them interpret and create graphs. (Schemata are frameworks located in long-term memory that organize people's understandings and theories about their environment. For more information about these frameworks, see II. Key Underpinnings: Schemata.) Students often lack graph-reading schemata; some students, for example, view graphs as literal pictures (Leinhardt, Zaslavsky, & Stein, 1990; McDermott, Rosenquist, & van Zee, 1987). When students are less skilled in graph reading, they have more difficulty identifying trends (Shah & Shellhammer, 1999). When a student does not automatically recognize the relationship encoded by a visual feature, a graph is harder to interpret (Cleveland, 1993; Kosslyn, 1994; Pinker, 1990; Shah & Carpenter, 1995; Shah, Mayer, & Hegarty, 1999).

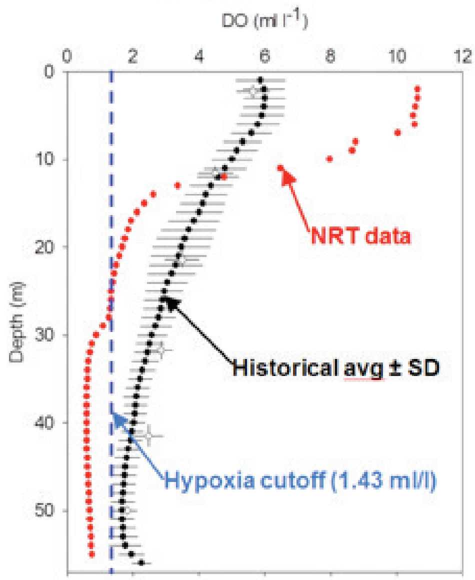


Figure 28. This graph of dissolved oxygen amount versus depth shows near-real time data (dissolved oxygen at different depths, in red), a historical average (dissolved oxygen, in black), and the hypoxia cutoff value (the minimal level of dissolved oxygen necessary to keep most sea animals alive, in blue). (Source: Modified Figure 4 from Grantham, B.A., F. Chan, K.J. Nielsen, D.S. Fox, J.A. Barth, A. Huyer, J. Lubchenco and B.A. Menge, 2004. Upwelling-driven nearshore hypoxia signals ecosystem and oceanographic changes in the northeast Pacific. Nature, 429, pp. 749-754.)

Students also often lack content-knowledge schemata that is related the to data on a graph. Research has shown that students are more likely to recognize patterns in graphs when they are familiar with the relationships among data variables (Freedman & Smith, 1996; Jennings, Amabile, & Ross, 1982; Shah, 1995; Shah & Shellhammer, 1999).

GUIDELINES

ENGAGE STUDENTS (GUIDELINES 40 AND 41)

Guideline 40: Pre-select data that are interesting to students

The graph comprehension process is difficult for students. To encourage them to persevere and overcome the difficulties they may encounter when they set out to explore data using graphs, pre-select data that will be familiar and interesting to students and that can answer questions that students find important.

Guideline 41: Make the interface design current and the graphs visually appealing

Most students are consumers of the Web, and they are more engaged in a Web-based task when the design appears current (W.Finzer, personal communication, June 29, 2011). It is also important to make graphs visually appealing whenever possible (Kosslyn, 2006).

PROVIDE INFORMATION ABOUT DATA SETS (GUIDELINES 42–44)

Guideline 42: Accompany data sets with conceptual organizers

As previously discussed in the CONSIDERATIONS for this section, the abstract nature of data sets creates difficulties for students. One way to make a data set less abstract is to connect it to a familiar process that incorporates the data. Ben-Zvi (2002) found that when students are given ample opportunity to learn about the phenomena that the data describe, their understanding of the data set improves. A conceptual organizer such as that shown in Figure 26, linked to a data set can also illustrate a relevant Earth process and help students make sense of the data and trends in a graph.

Guideline 43: Explain how the data were measured

To support students' use of expert data sets, links to explanations of how data were collected are especially important and should be provided. For example, Figure 26 illustrates the equipment used to collect dissolved oxygen data.

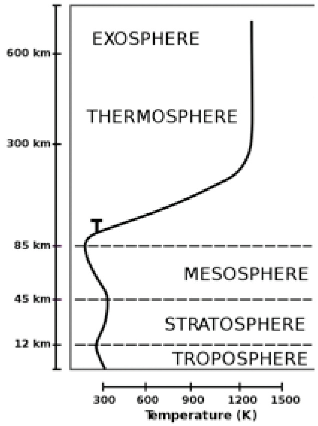


Figure 29. The boundaries highlight meaningful visual chunks related to atmospheric layers. (Source: Wikipedia, File:Atmosphere_with_Ionosphere.svg.png. http://en.wikipedia.org/w/index.php?title=File:Atmosphere_with_Ionosphere.svg&page=1)

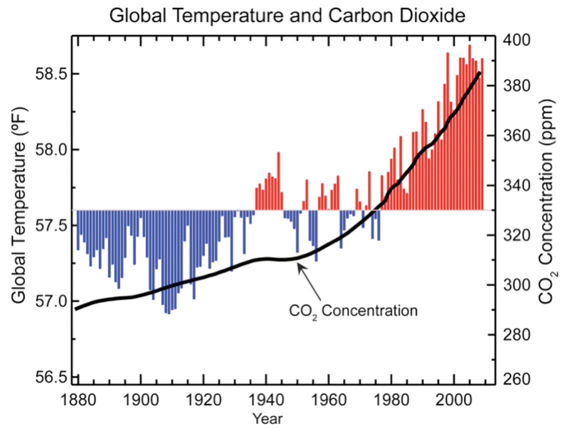


Figure 30. The colors in the graph highlight meaningful visual chunks related to anomaly, i.e., temperatures above and below the overall average. (Source: National Climate Data Center, National Oceanic and Atmospheric Administration. Retrieved from http://www.ncdc.noaa.gov/indicators/.)

Guideline 44: Describe important values and trends associated with each data set

Another way to make a data set less abstract is to include supplemental information that gives further details about the data. These supplements can be quantitative (e.g., historical average and benchmark values) or qualitative (e.g., verbal descriptions of historical trends and shifts in the data). Quantitative information can also be incorporated into graphs of the data set, as shown in Figure 28.

DEFINE AND REDUCE VISUAL CHUNKS (GUIDELINES 45–49)

Guideline 45: Use visual design to define visual chunks

When related parts of visual features are grouped by perceptual proximity, it is easier for students to identify visual chunks and recognize visual patterns. (For more information on grouping, see II. Key Underpinnings: Visual Perception and Processing.) Carpenter and Shah (1998) and Shah, Mayer, and Hegarty (1999) found that grouping of visual features positively influences viewers' spontaneous interpretations of graphs. It is important to understand that it is visual grouping rather than the type of graph, per se, that helps students recognize visual chunks. Color hues, boundaries, and placing related data in spatial proximity are ways to group visual chunks, which also minimizes extraneous cognitive load (Vekiri, 2002). Figure 29, a graph of altitude against temperature, uses boundaries to chunk the line into sections correlated to atmospheric layers.

Figure 30, a graph of temperature anomaly against year, uses color to chunk the line into sections that are less than average and greater than average (the "0" line is the average value of the entire data set).

Guideline 46: Provide a tool to draw a trend line

Multiple visual chunks of data make a graph difficult to interpret (Ratwani et al., 2008). Drawing one or more statistical trend lines can help students interpret a graph by reducing the complexity of the visual features and by showing trends that are hard to distinguish by eye. Figure 31 shows a trend line for a graph of precipitation percentage change against year in the United States.

A trend line can be drawn over a part of a graph. Figure 32 shows three trend lines for a graph of snow melt date (day of the year the snow's disappearance was visually observed) against year. The black trend line is for the entire graph, and the green and red trend lines are for parts of the graph. Although not shown in Figure 32, a trend line tool can also include slope calculations for each trend line.

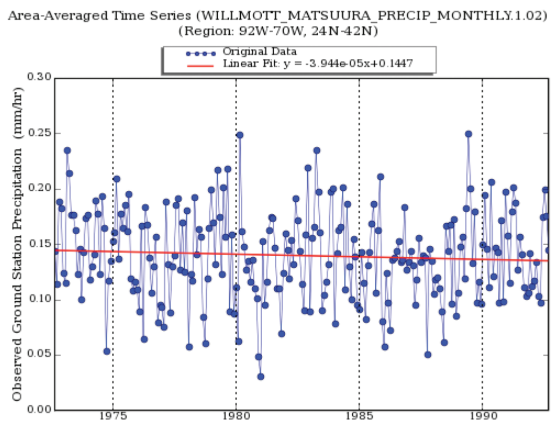


Figure 31. Adding a trend line makes the change in precipitation in this graph easier to see. (Source: Generated via Goddard Earth Sciences Data and Information Services Information Center, http://disc.sci.gsfc.nasa.gov/giovanni/overview/index.html.)

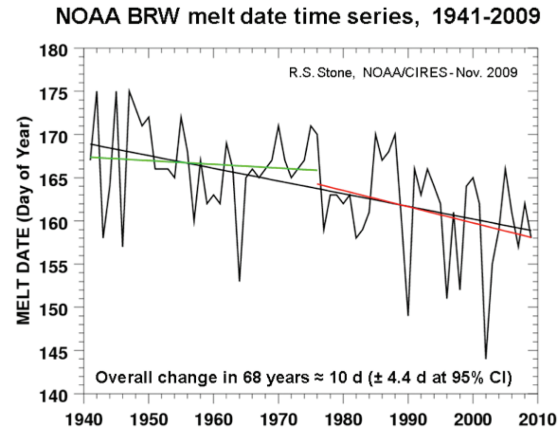


Figure 32. This graph shows the change in the snow melt date, based on visual observation by NOAA's Barrow observatory (BRW), between 1940 and 2010. The green and red lines show the change in slope between the first half and last half of the time series (Source: Earth System Research Laboratory Global Research Devision, National Oceanic and Atmospheric Administration. Retrieved from http://www.esrl.noaa.gov/gmd/grad/graphics/BRWmelt19412009.png.)

Guideline 47: Provide a tool to reduce the number of data points by averaging

Variability in large data sets can make visual chunks and trends difficult to see. The data sets can be reduced by using averaging. Figure 33, a grouped line graph of global surface temperature and sun's energy reaching top of the atmosphere versus time, shows collected and averaged data for both data sets.

Figure 33 shows the averaging of continuous data in a line graph. In contrast, Figure 34, a graph of the brightness of a star against time, shows the averaging of points in a scatter plot. The collected data are yellow and blue, and the averaged data are white.

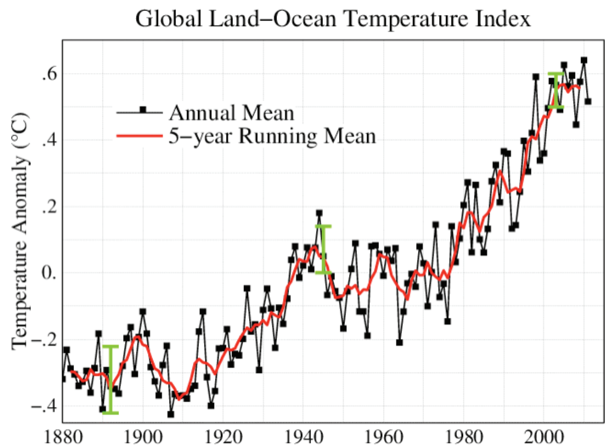


Figure 33. The trend of data can be made more clear by averaging, as shown by the red line. This line plot shows global mean land-ocean temperature index, 1880 to present, with the base period 1951-1980. The dotted black line is the annual mean and the solid red line is the five-year mean. The green bars show uncertainty estimates. (Source: Update of Figure 1A in J. Hansen, M. Sato, R. Ruedy, K. Lo, D.W. Lea and M. Medina-Elizade, 2006. Global temperature change. Proceedings of the National Academy of Sciences, 103, pp. 14208-14293. Retrieved from http://data.giss.nasa.gov/gistemp/graphs_v3/.)

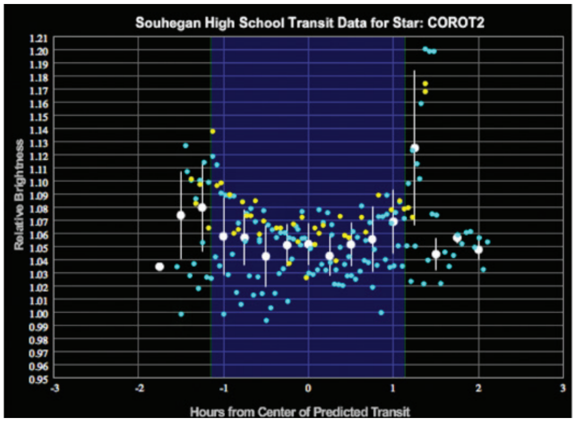


Figure 34. The trend of scatter plot data can be made more clear by averaging, as shown by the white points, which also include error bars. The blue-shaded area helps to focus students' attention on the time period during which the transit was predicted. (Source: Other Worlds / Other Earths, Harvard-Smithsonian Center for Astrophysics. Retrieved from http://www.cfa.harvard.edu/smgphp/otherworlds/index.php)

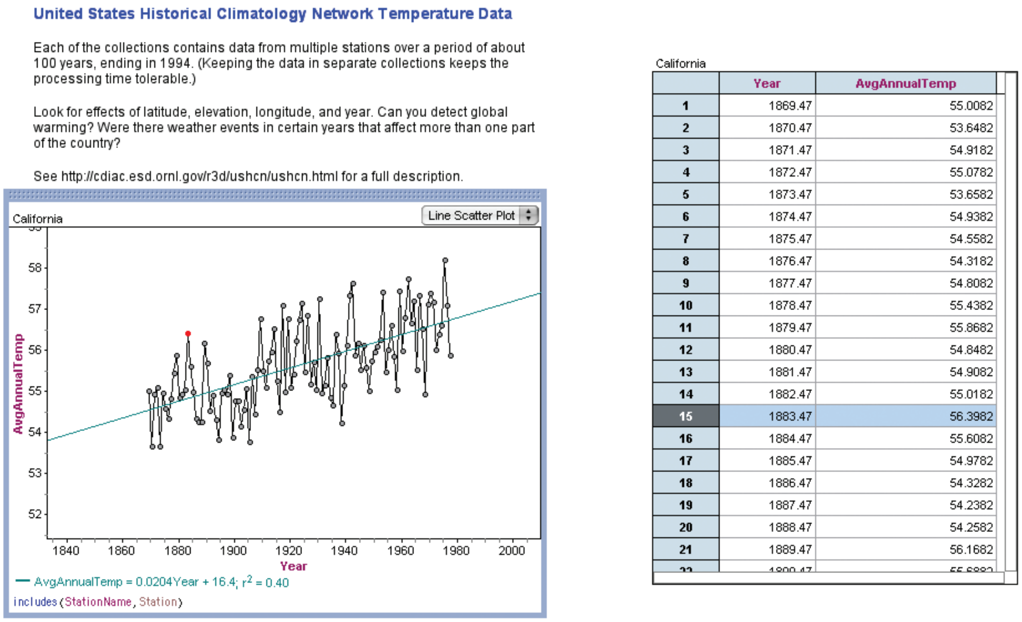


Figure 35. Visual chunks and textual information are integrated by use of color and direct labeling. (Source: Puget Sound Vital Signs, Puget Sound Partnership. Retrieved from http://www.psp.wa.gov/vitalsigns/images/herring_chart.png)

Guideline 48: Provide a tool to change the graph axes' scale and aspect ratio

Automatic scaling can obscure data details and affect the density of data points in scatter plots, making trends difficult to see and variability difficult to understand (Cleveland, 1993; Cleveland, Diaconis, & McGill, 1982; Lauer & Post, 1989). Providing a tool to change the scale and aspect ratio (the ratio of width to height of a graph) can support students in identifying visual chunks and trends and more clearly identifying patterns in variability.

Guideline 49: Provide a tool for removing data points

Some data points are the result of error, and removing these "outliers" can clarify the trend of the data. However, a removal should always be justified. Make sure that students are able to explain their reasoning when removing what they perceive to be erroneous data.

FACILITATE THE INTEGRATION OF VISUAL CHUNKS AND TEXTUAL INFORMATION (GUIDELINES 50 AND 51)

Guideline 50: Simplify the textual information

If textual information that is not needed for comprehension is included on a graph, extraneous cognitive load is increased. Simplify textual information by using easy-to-read fonts, removing unnecessary labels or markings, and identifying an axis or label with familiar symbols or icons.

Guideline 51: Use perceptual proximity to link visual chunks and textual information

Perceptual proximity facilitates integration by reducing cognitive load (Wickens & Carswell, 1995). Visual chunks and textual information that need to be linked in a graph should be close in perceptual proximity, either located near each other in space or connected via some other organizing feature, such as color hue. Direct labeling on a graph (rather than using legends) makes graph comprehension easier for students (Lewandowsky & Spence, 1989). The graph in Figure 35 integrates the visual chunks and textual information by directly labeling lines. The lines in the graph are labeled directly.

MAKE IT EASY TO EXAMINE AND EXTRACT QUANTITATIVE INFORMATION (GUIDELINES 52–56)

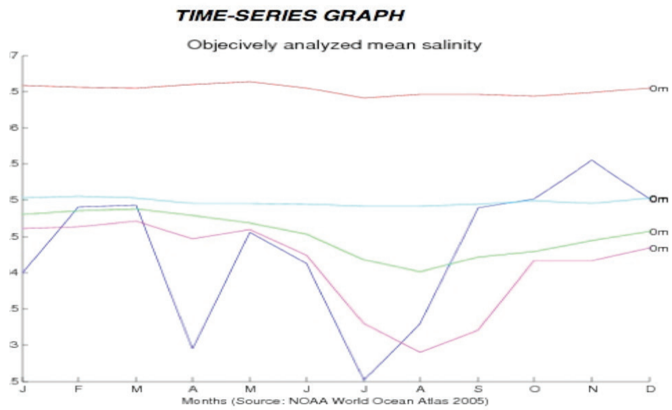
Guideline 52: Show data point values

Students often need to review individual data points before looking for visual chunks and trends in graphed data (Ben-Zvi, 2002). Tools that show point values include those that enable presentation of a data set table and graph simultaneously, those that allow a student to "mouse over" a point and see its value, and those that add grid lines to line graphs and bar graphs. Tables and graphs can be dynamically linked so that if a student clicks on a data point in one representation, that same data point is highlighted in the other representation. Figure 36 shows a scatter plot and corresponding data set table created using Fathom Dynamic Data™ software developed by William Finzer of KCP Technologies. When selected, points on the scatter plot are linked to values in the table using highlighting.

Guideline 53: Use visual design to show individual data points in a meaningful way

As described in the previous guideline, students often need to review individual data points. When students begin to work with large data sets, it is important that they transition from a focus on individual data to a focus on trends in the data set (Hammerman, 2009). To facilitate this transition, use visual design to show individual data points in a manner that also allows students to see trends in the larger data set. For example, the number of individual data points at one location in a scatter plot can be shown with partially transparent data points, so that color saturation is related to the number of points.

The number of individual data points within specific ranges on the x-axis or y-axis (or both) in a scatter plot can be shown with "bins," a feature that can help students identify trends in data. Figure 37 is a bin scatter plot of expenditures per student against average teacher salary, created using the TinkerPlots® Dynamic Data Exploration software.



Salinity (PSU)					
Month	(27.5N,89.3W) at 0m	(32.3N,61.9W) at 0m	(45.1N,44.2W) at 0m	(56.5N,31.8W) at 0m	(71.8N,13.3W) at 0m
1	34.001	36.587	34.799	35.033	34.616
2	34.907	36.555	34.854	35.056	34.641
3	34.933	36.545	34.883	35.033	34.718
4	32.951	36.597	34.783	34.958	34.473
5	34.554	36.643	34.683	34.954	34.595
6	34.131	36.545	34.53	34.937	34.244
7	32.52	36.407	34.181	34.91	33.302
8	33.296	36.46	34.009	34.921	32.897
9	34.887	36.457	34.21	34.936	33.203
10	35.012	36.438	34.29	34.989	34.16
11	35.554	36.483	34.448	34.954	34.165
12	35.024	36.551	34.572	35.031	34.347

Figure 36. This grouped line graph and data set tables show monthly average data of sea surface salinity at five different locations, with color-coded titles and lines. (Source: Jet Propulsion Laboratory, California Institute of Technology. Retrieved from http://aquarius.jpl.nasa.gov/AQUARIUS/index.jsp.)

Guideline 54: Take the accuracy of graphical-perception tasks into account

If it is important for students to automatically extract quantitative information from a particular graph, use the visual features that are most accurately perceived, as shown in Table 1, earlier in this section.

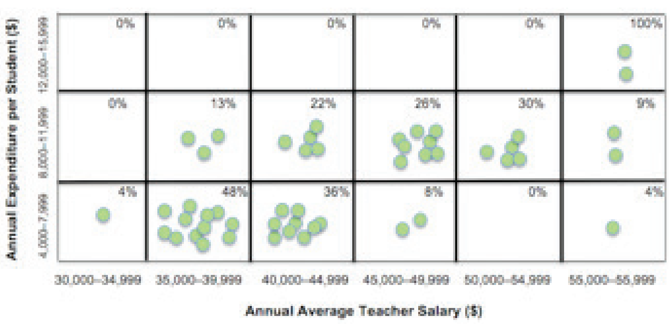


Figure 37. Placing data points in bins makes it easy to see both individual points and trends in the data. (Source: WolfWikis, North Carolina State University. Retrieved from http://wikis.lib.ncsu.edu/index.php/US_Schools_Group_C.)

Guideline 55: Use familiar quantities and units

Data from a scientific database are often labeled with abbreviations, quantities, and units that students have never seen, and the graphs created from those data sets will incorporate those labels. As much as possible, ensure

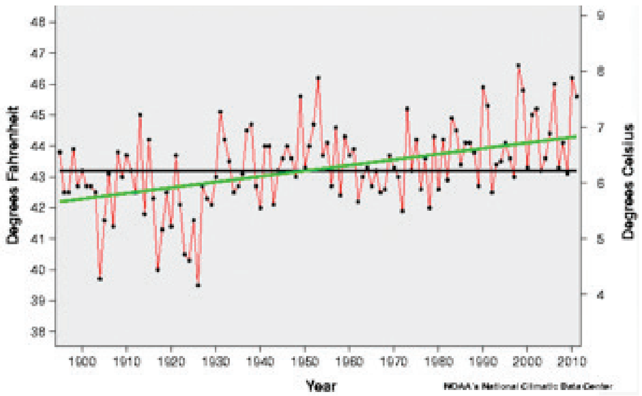


Figure 38. This graph shows temperature data for New Hampshire between 1895 and 2012. The black line is the average for the entire data set, the black dots and red line represent annual averages, and the green line shows the trend for the entire time period. (Source: Generated via National Climatic Data Center, National Oceanic and Atmospheric Administration, http://www.ncdc.noaa.gov/oa/climate/research/cag3/state.html.)

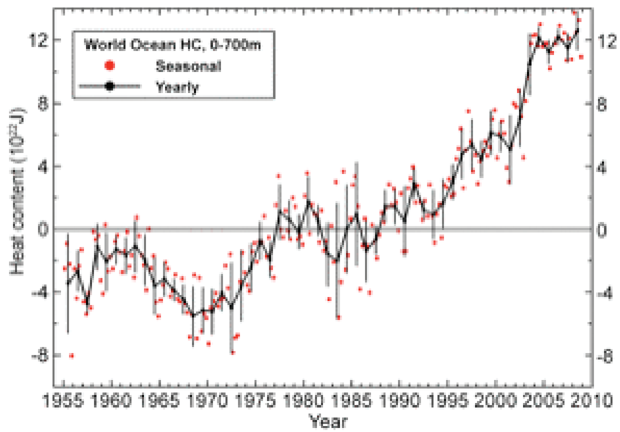


Figure 39. In an anomaly chart, the "0" value is the average for the entire data set. (Source: National Climatic Data Center, National Oceanic and Atmospheric Administration. Retrieved from http://www.ncdc.noaa.gov/indicators/.)

that quantities and units in data sets and graphs are written out and are familiar to students. If unfamiliar terms or abbreviations must be used, include an explanatory caption (Kosslyn, 2006).

Guideline 56: Provide a tool to show the measures of center

Measures of center, such as average, are important quantitative values associated with a data set and can be used to compare data sets (Konold & Pollatsek, 2002). To support students in working with measures of center, provide a tool that can calculate and show those measures in a graph. Figure 38 shows the average yearly temperature, the overall average temperature, and the overall trend line for temperature over a more than 100-year period in New Hampshire.

Another way to show average is to represent data in an anomaly chart, where measurements are shown as differences from the average value, which is set as 0. Figure 39, an anomaly chart, shows the upper ocean heat content between 1955 and 2010.

BRIDGE THE GAP FROM NOVICE TO EXPERT (GUIDELINES 57–60)

Guideline 57: Provide graph-reading supports

The ability to comprehend graphs depends on the reader's graph schemata, or general knowledge about graphs, and many students have limited knowledge in this area. Diagrams that have labels or notes explaining visual features and graph details support students' comprehension (Vekiri, 2002). It is also important to make graph reading metacogni-

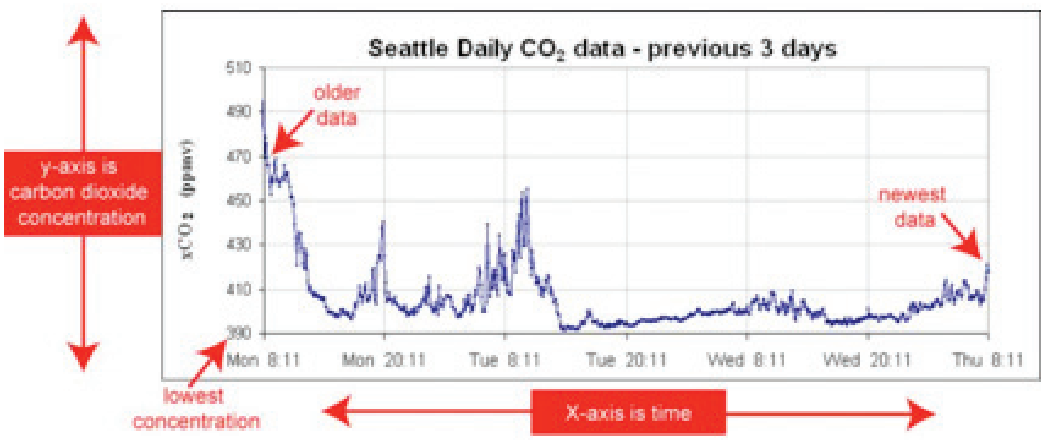


Figure 40. This graph is annotated with notes that highlight the axes' labels and explain how values change along each axis. (Source: Pacific Marine and Environmental Laboratory, National Oceanic and Atmospheric Administration. Retrieved from http://pmel.noaa.gov/co2/files/sn_3day_plot.jpg.)

tive by explaining to students that graph reading is an interpretation-and-evaluation task as opposed to a mere fact-retrieval task (Shah & Hoeffner, 2002). Examples of linked graph-reading supports are shown in Figures 38 and 39.

Guideline 58: Provide content-knowledge supports

An education interface will often focus instruction on a limited set of phenomena. Wherever it seems appropriate, provide links to content-knowledge supports, such as diagrams, text, and glossaries. The diagram in Figure 41 explains the acidification of the ocean in detail.

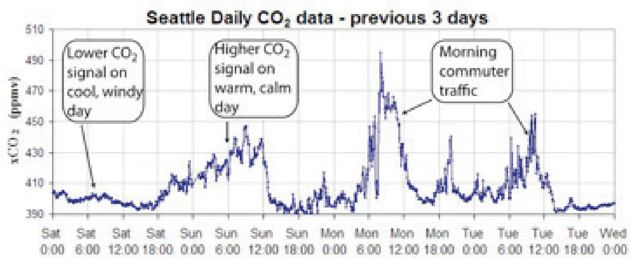


Figure 41. This graph is annotated with notes that explain common trends seen in CO2 emissions data. (Source: Pacific Marine and Environmental Laboratory, National Oceanic and Atmospheric Administration. Retrieved from http://pmel.noaa.gov/co2/files/sn_3day_plot.jpg.)

Guideline 59: Provide annotation and note-taking tools

In an education interface, students may create multiple representations of graphs and/or linked data tables and graphs. It is important to provide a way for students to annotate their work and take notes so they can keep track of conclusions, observations, and questions about various representations. Students can also use these notes for discussion and reflection, practices that experts carry out.

Guideline 60: Provide metacognitive instructions for graph creation and interpretation

Kramarski (2004) found that students' misconceptions were lessened by metacognitive instructions asking them to reflect on their problem-solving process. These instructions guided students to consider the task, look for relevant information, and consider what they already know that relates to the current task.

FACILITATE CREATION OF MULTIPLE VIEWS AND COMPARISONS OF MULTIPLE GRAPHS (GUIDELINES 61–63)

Guideline 61: Provide a tool to add and subtract data subsets

Data subsets can be added to and subtracted from graphs to build an understanding of the relationships among data. Typically, trends in scatter plots become clearer when data are added. Figure 43 shows the use of a tool that allows students to add and subtract data subsets. In this example, the two data subsets correspond to collections by students from different classrooms.

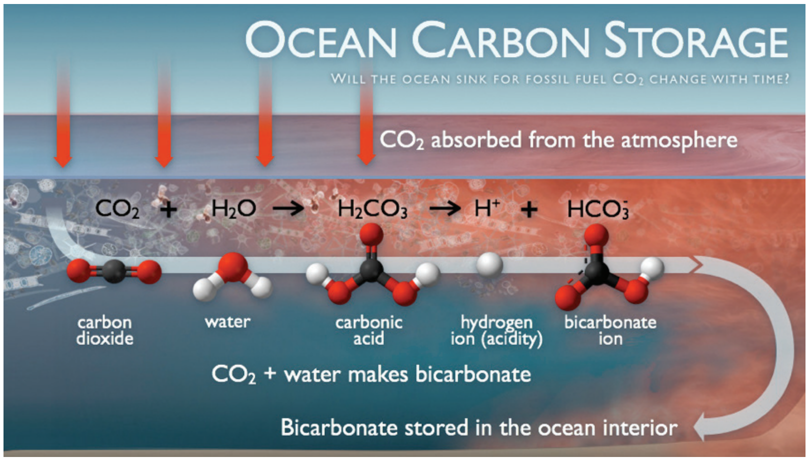


Figure 41. This diagram, an example of a content-knowledge support, explains how absorbed CO2 acidifies the ocean. (Source: Pacific Marine and Environmental Laboratory, National Oceanic and Atmospheric Administration. Retrieved from http://www.pmel.noaa.gov/co2/story/Ocean+Carbon+Storage.)

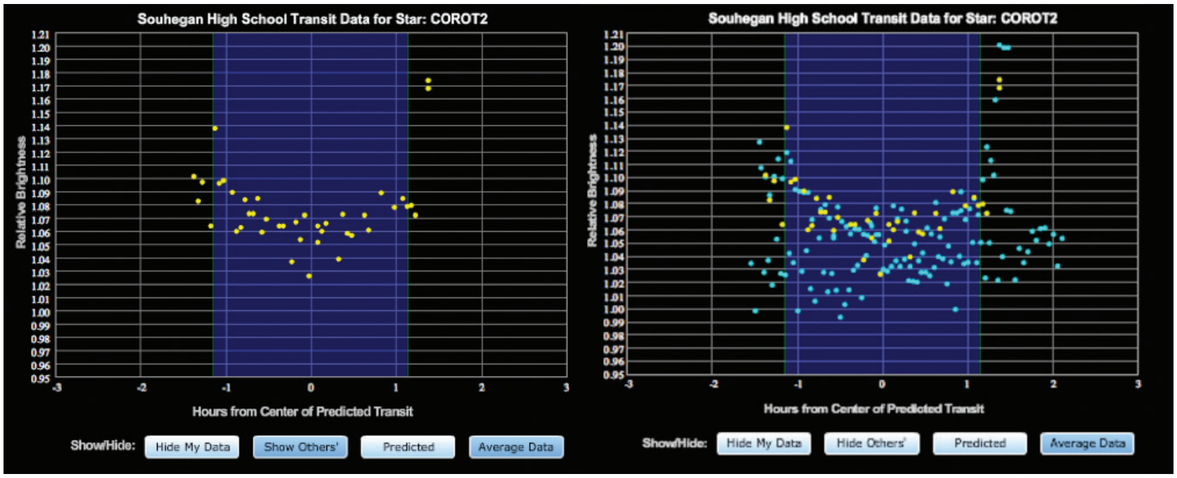


Figure 43. Two data subsets—color-coded yellow and turquoise—are added to a graph. (Source: Other Worlds / Other Earth, Harvard-Smithsonian Center for Astrophysics. Retrieved from http://www.cfa.harvard.edu/smgphp/otherworlds/index.php)

Guideline 62: Provide a tool to create multiple graphs of the same data set

Viewing multiple graphs of the same data allows students to gain content knowledge, because they see different information from the data highlighted, and to gain familiarity with graphs in general (Shah & Hoeffner, 2002). This tool can facilitate comparison by dynamically linking data in the graphs, automatically sizing and positioning the different graphs side by side, and standardizing text, labels, colors, and features.

Guideline 63: Provide tools to compare graphs of different data

Students may want to compare graphs of different data sets. For example, a student might want to compare graphs of the same measurements taken at different locations, or graphs of different data that are related to the same Earth process. As in the previous guideline, a tool that allows them to compare different graphs can facilitate comparison by dynamically linking data in the graphs, automatically sizing and positioning different graphs side by side, and standardizing labels, colors, and features.

If the data sets share a common x-axis, comparison of multiple graphs is easier when bar charts and line graphs are presented as a group and include a common baseline (Lewandowsky & Spence, 1989). Otherwise, the graphs should be shown side by side. Kosslyn (2006) recommends additional strategies for easier comparison of multiple graphs:

- When comparing multiple data sets in a grouped line graph or a grouped bar graph, include no more than four different sets.
- Label all lines or bars that are to be compared. However, if one line or bar is already highlighted in the title, it doesn't need a label.
- If a legend is necessary due to space limitations, place it in the top right corner.
- If two graphs share a common y-axis, line up the graphs side by side and include only one y-axis on the far left.
- Give students the option to assign one data set as the primary data set so that its representation can be emphasized.
- Another way to emphasize a data set is to make the visual features that represent it darker or thicker.

REFERENCES

Ben-Zvi, D. (2002). Seventh grade students' sense making of data and data representations. In B. Phillips (Ed.), *Proceedings of the Sixth Annual Conference on Teaching of Statistics, Cape Town, South Africa*. Voorburg, The Netherlands: International Statistical Institute.

Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied, 4*(2), 75-100.

Cleveland, W. S. (1993). *Visualizing data*. Summit, NJ: Hobart Press.

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association, 79*(387), 531-554.

Cleveland, W. S., Diaconis P., & McGill, R. (1982). Variables on scatterplots look more highly correlated when the scales are increased. *Science, 216,* 1138-1140.

Freedman, E. G., & Smith, L. D. (1996). The role of data and theory in covariation assessment: Implications for the theory-ladenness of observation. *Journal of Mind and Behavior, 17*, 321-343.

Gillian, D. J., & Lewis, R. (1994). A componential model of human interaction with graphs: 1. Linear regression modeling. *Human Factors, 36*(3), 419-440.

Hammerman, J. (2009, August 16-22). *Informal inferential reasoning about large scientific data sets.* Paper presented at the International Statistical Institute Conference, Durban, South Africa. Retrieved from http://www.statssa.gov.za/isi2009/ScientificProgramme/IPMS/0376.pdf

Hammerman, J., & Rubin, A. (2004). Strategies for managing statistical complexity with new software tools. *Statistics Education Research Journal, 3*(2), 17-41.

Jennings, D. L., Amabile, T. M., & Ross, L. (1982). Informal covariation assessment: Data-based versus theory-based judgments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 211–230). New York, NY: Cambridge University Press.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education, 33*(4), 259-289.

Konold, C., Higgins, T., Russell, S. J., & Khalil, K. (2004). *Data seen through different lenses.* Unpublished manuscript. Retrieved from http://www.srri.umass.edu/publications/konold-2004dst

Kosslyn, S. M. (1994). *Elements of graph design.* New York, NY: Freeman.

Kosslyn, S. M. (2006). *Graph design for the eye and mind.* New York, NY: Oxford University Press.

Kramarski, B. (2004). Making sense of graphs: Does metacognitive instruction make a difference on students' mathematical conceptions and alternative conceptions? *Learning and Instruction, 14,* 593-619.

Latour, B. (1990). Visualisation and cognition: Drawing things together. In M. Lynch & S. Woolgar (Eds.), *Representation in Scientific Practice* (pp. 19-68). Cambridge, MA: MIT Press.

Lauer, T. W., & Post, G. V. (1989). Density in scatterplots and the estimation of correlation. *Behaviour and Information Technology, 8*, 235-244.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research, 60*, 1-64.

Lewandowsky, S., & Spence, I. (1989). *The perception of statistical graphs. Sociological Methods and Research, 18,* 200–242.

McDermott, L., Rosenquist, M., & van Zee, E. (1987). Students' difficulties in connecting graphs and physics: Example from kinematics. *American Journal of Physics, 55*(6), 503-513.

Mokros, J., & Russell, S. J. (1995). Children's concepts of average and representativeness. *Journal for Research in Mathematics Education, 26*(1), 20-39.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73-126). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ratwani, R. M., Trafton, J. G., & Boehm-Davis, D. A. (2008). Thinking graphically: Connecting vision and cognition during graph comprehension. *Journal of Experimental Psychology: Applied, 14*(1), 36–49.

Shah, P. (1995). *Cognitive processes in graph comprehension.* Unpublished doctoral dissertation. Carnegie Mellon University, Pittsburgh, PA.

Shah, P., & Carpenter, P. A. (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General, 124,* 43-62.

Shah, P., & Hoeffner, J. (2002). Review of Graph Comprehension Research: Implications for Instruction. *Educational Psychology Review, 14*(1), 47-69.

Shah, P., Mayer, R. E., & Hegarty, M. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology, 91*, 690-702.

Shah, P., & Shellhammer, D. (1999, July 9-11). *The Role of Domain Knowledge and Graph Reading Skills in Graph Comprehension.* Presented at the 1999 Meeting of the Society for Applied Research in Memory and Cognition, Boulder, CO.

Tufte, E. R. (1983). *The Visual Display of Quantitative Information.* Cheshire, CT: Graphics Press.

Vekiri, I. (2002). What Is the Value of Graphical Displays in Learning? *Educational Psychology Review, 14*(3), 261-312. Retrieved from http://hdl.handle.net/2027.42/44453

Wickens, C. D., & Carswell, C.M. (1995). The proximity compatibility principle: Its psychological foundations and its relevance to display design. *Human Factors, 37*, 473-494.

Animations

CONSIDERATIONS

WHAT ARE WE TALKING ABOUT?

Animated visualizations include one or more features that exhibit change over time. This change can take on a variety of forms, including transformation (changes to an object's attributes, such as size or color), translation (changes to location in 2D or 3D space), and transitions (appearance or disappearance from a visualization) (Lowe, 2003).

There are multiple ways that animated visualizations might be integrated into an educational interface for scientific data, three of which are the primary focus of this section:

- **Concept-driven:** A concept-driven visualization is developed to illustrate a concept or theory and is not directly tied to empirical data (Clark & Wiebe, 2000). In an interface, dynamic diagrams of this type might deliver supplemental or supportive information that provides context for a data visualization. Many of the scientific phenomena that students are likely to investigate using Earth science data are quite complex. Explanatory visuals, including animations, may be an important part of the supports provided to help scaffold user understanding of these complex systems. In addition, explanatory animated diagrams might help users connect abstract data to its origins in the physical world. For example, a student might view a short animation showing how a glider moves through the water at various depths, latitudes, and longitudes, remotely collecting data.
- **Data-driven:** Data-driven visualizations use empirical or model data and represent variation in data values using graphical elements (Clark & Wiebe, 2000). In the previous sections on geo-referenced data visualizations and graphs, we saw that properties such as color, size, and shape can be used to communicate information about data values. In animations, motion can also be used to indicate variation in a data set.
- **Transitional:** An interface might feature animated transitions as users move between visualization formats (e.g., displaying a data set on a bar graph and then a pie chart), between data sets (e.g., displaying data sets from two different time periods on the same scatterplot), and as they manipulate visualization parameters, such as scale or point of view. While there is some empirical support for the latter in the context of linking multiple data displays, more research is needed on the effective use of other animated transitions between data displays.

Two other examples of dynamic representations are animated simulations and video footage; however, while many of the guidelines below may apply to these visualizations, they are not a major focus of this section.

WHAT SHOULD THE INTERFACE DESIGNER CONSIDER?

The jury is still out on whether animations can facilitate learning. As technology becomes more ubiquitous in homes and classrooms, the use of educational multimedia is on the rise, and dynamic multimedia, such as animations, appeal to many as a method for engaging end users and communicating information in digital environments. However, research on the effective design and use of animations in educational contexts is far from conclusive (Hegarty, 2004; Hoffler & Leutner, 2007; Tversky, Morrison, & Betrancourt, 2002). While educational animations provide promising tools for enhancing learning, the potential pitfalls are numerous, and a number of studies suggest that animated visualizations are no more effective than their static counterparts (and in some cases are even less so!).

According to principles of multimedia design, the structure and content of an external representation should correspond to the structure and content of the material it represents (Betrancourt, 2005). In theory, then, animated visualizations are a natural choice for illustrating changes that occur over time. In the context of exploring data, animations show the potential to enhance user understanding, for example, by facilitating comparisons (Nakakoji, Takashima, & Yamamoto, 2001) or drawing users' attention to specific aspects of a data set (Gordin, Polman, & Pea, 1994). However, research also indicates that in order for students to reap these kinds of benefits, animations must be designed to address the challenges imposed by fleeting sources of information. It is important for interface designers to understand the potential difficulties that users might face and to design strategies that will support the effective use of animated visualizations, in appropriate contexts.

Novices have difficulty recognizing and remembering important changes. A variety of factors make novice users particularly susceptible to missing important elements in an animated visualization (Rieber, 1990, 2000). All users may experience "change blindness" due to natural saccades or eye shifts (Grimes, 1996), pauses—even brief ones—between frames (Pashler, 1988), and the complexity of the content (Pylyshyn & Storm, 1988). However, without the ample experience with data visualizations and the well-developed schemata enjoyed by scientists who know what to look for, novices face additional risks. Their eyes and attention are naturally drawn to the changes that perceptually "pop out," such as motion, which means that they often miss non-dynamic elements (MacEachren, 2004; Nossum, 2010) and/or more subtle changes (Lowe, 2003). In addition, evidence suggests that novices tend to look for cause-effect relationships or causal agents in animations, whereas experts are likely more attuned to more complex relationships and are less quick to assign cause-effect relationships erroneously (Narayanan & Hegarty, 2000). Finally, novice users' existing misconceptions can direct their attention inappropriately as they look for elements or relationships based on their inaccurate or incomplete understandings (Plass, Homer, & Hayward, 2009; White, 1984). It is important that interface designers consider that change-blindness resulting from any of these factors can lead to the development of inaccurate or incomplete understandings.

Animated maps are particularly challenging for novice users. Novice users who have trouble reading static maps can find it even more difficult when these representations are animated—and many of the challenges that novices face with animations in general are exacerbated when they must grapple with an animated map. Maps require a lot of searching and scanning, even in their static forms and they typically feature more subtle transformations (e.g., changes to hue or size) to indicate changes in data values, making users increasingly susceptible to change-blindness (Goldsberry & Battersby, 2009). Also, animated maps usually require users to refer back to both timelines and legends, forcing them to divide their attention and increasing their cognitive load (Opach & Nossum, 2011).

GUIDELINES

MINIMIZE THE EXTRANEOUS COGNITIVE LOAD AND MITIGATE THE INTRINSIC COGNITIVE LOAD ASSOCIATED WITH ANIMATIONS (GUIDELINES 64–69)

While a number of demands are placed on users' cognitive resources as they perceive, process, and interpret any data visualization, animations will inherently impose additional extraneous cognitive load (i.e., additional demands due to the format of the visualization itself), as users must perceive and hold a constantly changing source of information in their working memory (Ayres & Paas, 2007; Spanjers, Wouters, van Gog, & van Merriënboer, 2011). If sufficient cognitive resources are not available, users run the risk of developing inaccurate or incomplete understandings of the material. Therefore, it is critical that animated visualizations be designed to alleviate the demands associated with transient information and to avoid design elements that add to processing demands without enhancing learning.

Guideline 64: Eliminate unhelpful redundancies

For some users, the inclusion of redundant information imposes additional extraneous cognitive load, making it more difficult to extract important features or information (Hasler, Kersten, & Sweller, 2007). Redundancies can take on a variety of forms, for example, integrating labels to accompany iconic elements or including captions describing features and relationships that are self-evident in the visualization. Basically, designers should avoid duplicating information if one of the sources is intelligible on its own, as the user's prior knowledge will come into play when making that determination. Those with low prior knowledge may not be able to make sense of a visualization without supportive labeling or explanatory text, and in these cases redundancies can be quite helpful. However, for users who are knowledgeable or experienced in a content area, the redundant information requires additional processing without enhancing their understanding and may inhibit their ability to attend to other important features in the visualization. In order to design effective animations for an audience with a range of existing knowledge, it may be necessary to include redundant labels and supporting explanations as optional features that can be turned on or off as appropriate for individual or groups of users.

It should be noted that redundancies, whether helpful or unhelpful, do not seem to be a significant factor when the animation is sufficiently "easy" and does not approach the limits of the user's working memory (Mayer & Moreno, 2003).

Guideline 65: Maximize users' available working memory resources by presenting verbal explanations as audio narrations

Coordinating visual and verbal representations is thought to help novices build more coherent and complete schemata, based on the premise that each moves through unique processing channels and results in different forms of mental representations that can both reference and enhance each other (Mayer & Anderson, 1991; Paivio, 1986; Rieber, 1990). However, trying to integrate written verbal explanations into an animation will most likely lead to a situation that overburdens users' working memory. Users must either (1) take their eyes off important changes as they're happening in order to read the text, or (2) try to hold the visual or verbal information in their working memory if the two are presented one after the other.

The literature offers a solution. Research on learning from animated multimedia supports the theory that visual and auditory stores of working memory are at least partially independent, and that one way to avoid overburdening users' visual working memory resources and dividing their attention is to present verbal explanations and cues as narration rather than as accompanying text (Mayer & Moreno, 2003; Vekiri, 2002). This strategy might also be applied to legends in animated maps in the form of audible legends or audible supports to visual legends (Kraak, Edsall, & MacEachren, 1997).

Designers should bear in mind that narration is a tool ideally suited for material that presents significant intrinsic cognitive load for users. If narrated information is redundant and unnecessary, given the user's prior knowledge, it can still result in extraneous processing and distract from other important visualization details. Designers should also avoid any temptation to introduce audio elements intended to entertain or increase appeal, as there is no evidence that these "bells and whistles" actually improve understanding. In fact, they can actually distract users from the material that is of importance (Moreno & Mayer, 2000).

Guideline 66: Keep users' attention focused by visually and temporally integrating related information

Particularly for learners with lower prior knowledge, it is important to provide both graphic and verbal information and to ensure that this information is integrated temporally as well as spatially (Plass et al., 2009; Vekiri, 2002). For example, the designer can add labels or explanations next to the animated object, show two animations with related information at the same time or back to back, or include an auditory narration over the animation (Mayer, 1997).

Guideline 67: Use schematized depictions rather than overly realistic or symbolic ones

Research on effective animation design and research comparing animations to their static counterparts suggests that there can be disadvantages to using more realistic-looking features in a dynamic visualization versus using simplified or schematized versions of the same features (Hoffler & Leutner, 2007). While realistic depictions may facilitate recognition, the increased level of detail may also direct attention away from more critical aspects and add to processing demands. Schematizing, or simplifying visualizations to only include essential elements, allows information to be presented in an easier-to-perceive way (Imhof, Scheiter, & Gerjets, 2009). Optimally designed animations should include enough realistic detail to allow for quick and accurate mapping to users' existing schemata, while minimizing the search and perceptual processing necessary to identify critical features.

Designers should also avoid the temptation to include too many symbolic icons. While they may communicate more information to those with more prior knowledge and automated schemata, for novices, symbolic elements may cause confusion or require more working memory resources during recognition (Plass et al., 2009).

Guideline 68: Give users control over the progression and pace of animations

Allowing users some form of control over an animation's pace shows promise as a strategy for making animations feasible as educational tools. Whether users control the speed at which they progress through pre-identified segments or whether they have total control over the pace of an animation (using start/stop/pause buttons), interaction around an animation's progression seems to provide a significant advantage by helping to optimize cognitive load (i.e., reducing extraneous cognitive load and increasing germane cognitive load) (Betrancourt, 2005; Cook, 2006; Harrower & Fabrikant, 2008; Hasler et al., 2007; Kraak et al., 1997; Mayer & Chandler, 2001; Swaak & De Jong, 2001; Tabbers, Martens, & van Merrienboer, 2004; Velez, Silver, & Tremaine, 2005). Extraneous cognitive load may be reduced by giving users an opportunity to stop and process what they have seen at strategic points in the animation rather than processing it in its entirety at the end. For users with full control over the length of the segments and their pace going through the material, germane cognitive load may also be increased. Evidence shows that both map users and graph users want to have some control over animated versions of these visualizations (Koussoulakou & Kraak, 1992; Monmonier & Gluck, 1994; Nakakoji et al., 2001).

While there is considerable evidence in support of this strategy, it seems to be most effective when used with content that imposes a high level of intrinsic cognitive load (i.e., is challenging) for the user (Hasler et al., 2007). There also appears to be a fine line between increasing germane cognitive load and imposing extraneous cognitive load by having users engage with animation controls. For example, users without enough metacognitive abilities or who are unfamiliar with the controls might spend too many working memory resources trying to figure out the appropriate times to stop or how to use the controls (Chandler, 2004; Cook, 2006; Lowe, 2004). Finally, more research is needed on the optimal number of segments, the most helpful types of controls, and the overall speed of the animation (Hasler et al., 2007).

Guideline 69: Cue users' attention to important features and patterns

In order for an animated visualization to be effective, users must accurately attend to, perceive, and understand its critical components. Perception requires attention, either conscious or due to the biases of our visual system, and unlike their static counterparts, animations don't afford users the luxury of taking their time in the search for critical features and patterns (Goldsberry & Battersby, 2009; Simons & Chabris, 1999). Prior knowledge, existing misconceptions, and previous experience learning from animations can influence how users perceive and attend to details (Plass et al., 2009; White, 1984). Novices with limited knowledge and experience (and perhaps a few misunderstandings) in a content area have particular difficulty knowing how to attend to an animation's relevant cues or details (Rieber, 1990, 2000). Evidence indicates that user comprehension is increased when users are provided with cues that help focus their attention on important aspects, likely because users spend fewer cognitive resources searching for important features and because less noticable changes (e.g., changes in color or the disappearance of data points) become less susceptible to change-blindness (Rieber 2000; Vekiri, 2002). Interface designers should make use of the cross-cutting guidelines for drawing user attention to important features and patterns in order to support users as they work with animated data visualizations.

USE ANIMATIONS JUDICIOUSLY AND IN THE APPROPRIATE CONTEXTS (GUIDELINES 70–73)

As we've seen, while animations should be effective in depicting dynamic, time-based data, humans (particularly novice users with relatively less domain knowledge) often have difficulty learning from them. (Goldsberry & Battersby, 2009; Tversky et al., 2002). As a first step, before implementing any of the design strategies above, interface designers should make evidence-informed decisions about whether an animated visualization is the most effective one, based on the data to be visualized, the end users, and the likely tasks for which the data will be used.

Guideline 70: Only use animations to depict dynamic, time-based data—avoid introducing animated material solely for entertainment value or to increase engagement

Animations are most effective when the user needs to visualize continuous change over time and space. Rieber (1990) argues that animated diagrams might provide an advantage over static displays in tasks involving learning about dynamic phenomena because they depict motion and trajectory more effectively. Animations work less well when used to depict discrete change over time or when time is not a significant variable.

There can also be a tendency in educational multimedia to include features that make learning "fun," such as music or decorative movement, following the logic that an entertained user will be more motivated and engaged with the content. However, Harp and Mayer (1997) provide evidence to the contrary, finding that animated features such as these actually distract users from what they should be attending to. In general, animations have compared favorably to static visualizations when they are representational rather than cosmetic—that is, when the dynamic features are used to explicitly present the content to be considered (i.e., in our context, data and data-related phenomena) (Goldsberry & Battersby, 2009; Tversky et al., 2002).

Guideline 71: Use animations for specified tasks, and avoid using them for unstructured exploratory tasks

Animated visualizations, particularly animated maps, may not be well-suited for more open-ended, exploratory tasks. Part of effectively adjusting the cognitive load imposed by animations is minimizing the amount of non-essential, "extra" information that might distract the user. However, what is non-essential for one task or question may be quite pertinent in another context. It is much easier to implement guidelines intended to mitigate extraneous cognitive load with specified tasks in mind. In addition, novice learners often struggle with exploratory learning without sufficient support and guidance (Plass et al., 2009), and it is difficult to cue user attention to critical content and change processes when they're engaged in exploratory tasks using dynamic maps (Goldsberry & Battersby, 2009).

Guideline 72: Use animations to explicitly link representations

The dynamic linking of representations—so that one varies in response to an action that the user carries out with another—can facilitate students' work with multiple data representations. As Ainsworth (2006) explains:

> *Learners act on one representation and see the results of those actions in another. . . . Dynamic linking of representations is assumed to reduce the cognitive load upon the student—as the computer performs translation activities, students are freed to concentrate upon their actions on representations and their consequences in other representations. (p. 15)*

Linking representations reduces cognitive load by reducing the time and energy that students must spend searching for related features (Cook, 2006).

Data tables and graphs can be dynamically linked by highlighting points on the graph where data in the table are selected and vice versa, reinforcing students' understanding of the data represented in the graph (W. Finzer, personal communication, [June 29, 2011; Shah & Hoeffner, 2002). For example, the Fathom Dynamic Data™ software created by William Finzer of KCP Technologies allows students to move curves on a graph and then see how the equations defining those curves change, as shown in Figure 44.

Maps and data tables can also be dynamically linked. The Climate Visualizer software (Gordin, Polman, & Pea, 1994) allows students to select a location from a table and see it highlighted on a map. Dynamic linking can also help students stay oriented as they create a zoom-in view of an area by showing the region in the current view simultaneously in an associated small-scale map (Plaisant, Carr, & Shneiderman, 1995).

However, when deciding whether to dynamically link representations, keep in mind that for some more capable and experienced users, dynamic linking may be unnecessary and in fact may discourage users from making connections for themselves and thereby constructing deeper understandings (Ainsworth, 2006).

Guideline 73: Consider the relative strengths of and opportunities provided by static representations

While the research reviewed for this KSR suggests several potential benefits for the strategic use of well-designed animations in a data visualization interface, it also clearly indicates that this is an area not without challenges and in need of additional research. Depending on the nature of the data, the likely nature of he tasks, and the primary target audience of the interface, it may sometimes make more sense to take advantage of the relative strengths offered by static geospatial representations, graphs, and images (Tversky et al., 2002).

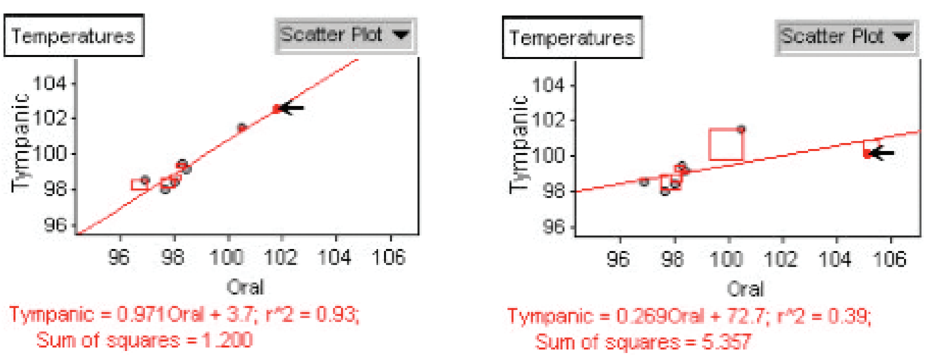


Figure 44. The Fathom software allows students to find the best-fit curve by minimizing the visual and numerical sum of squares in a graph. (Source: William Finzer, Fathom Dynamic Data Software, 2006.)

REFERENCES

Ainsworth, S. E (2006). DeFT: A conceptual framework for learning with multiple representations. *Learning and Instruction, 16*(3), 183-198.

Ayres, P., & Paas, F. (2007). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology, 21,* 695-700. doi:10.1002/acp.1343

Betrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 287-296). Cambridge, England: Cambridge University Press.

Chandler, P. (2004). The crucial role of cognitive processes in the design of dynamic visualizations. *Learning and Instruction, 14,* 353-357.

Clark, A., & Wiebe, E. (2000). Scientific visualization for secondary and post-secondary schools. *Journal of Technology Studies, 26,* 24-32.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education, 90*(6), 1073-1091.

Finzer, W. (2006). *Fathom Dynamic Data Software.* Emeryville, CA: Key Curriculum Press.

Goldsberry, K., & Battersby, S. E. (2009). *Issues of change detection in animated choropleth maps. Cartographica, 44*(3), 201-215.

Gordin, D. N., Polman, J. L., & Pea, R. D. (1994). The Climate Visualizer: Sense-making through scientific visualization. *Journal of Science Education and Technology, 3*, 203-226.

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Perception.* Vancouver Studies in Cognitive Science (vol. 5). New York, NY: Oxford University Press.

Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology, 89*, 92-102.

Harrower, M., & Fabrikant, S. (2008). The role of map animation in geographic visualization. In M. Dodge, M. Turner, & M. McDerby (Eds.), *Geographic Visualization* (pp. 49-66). Hoboken, NJ: Wiley and Sons.

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology, 21,* 713-729.

Hegarty, M. (2004). Dynamic visualizations and learning: Getting to the difficult questions. *Learning and Instruction, 14,* 343-351.

Höffler, T., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17,* 722-738.

Imhof, B., Scheiter, K., & Gerjets, P. (2009). Realism in dynamic, static-sequential, and static-simultaneous visualizations during knowledge acquisition on locomotion patterns. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2962-2967). Austin, TX: Cognitive Science Society.

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

Kraak, M. J., Edsall, R. M., & MacEachren, A. M. (1997). Cartographic animation and legends for temporal maps: Exploration and/or interaction. In ICC 1997: *Proceedings of the 18th ICA International Cartographic Conference* (vol. 1). Stockholm, Sweden: International Cartographic Association.

Lowe, R. K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction, 13,* 247-262.

Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction, 14,* 257-274.

MacEachren, A. M. (2004). *How Maps Work: Representation, Visualization, and Design.* New York, NY: The Guilford Press.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32,* 1-19.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of dual-coding hypothesis. *Journal of Educational Psychology, 83,* 484-490.

Mayer, R., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93,* 390-397.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38,* 43-52.

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

Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology, 92*(1), 117-125.

Nakakoji, K., Takashima, A., & Yamamoto, Y. (2001). Cognitive Effects of Animated Visualization in Exploratory Visual Data Analysis. *InfoVis '01: Proceedings of the IEEE Symposium on Information Visualisation* (pp. 77-84). Washington, DC: IEEE Computing Society.

Narayanan, N. H., & Hegarty, M. (2000). Communicating dynamic behaviors: Are interactive multimedia presentations better than static mixed-mode presentations? In *Proceedings of the First International Conference on the Theory and Application of Diagrams* (pp. 178-193). Berlin, Germany: Springer-Verlag.

Nossum, A. S. (2010, September 15-17). *Semistatic Animations—Integrating Past, Present and Future in Map Animations.* Presented at GIScience 2010, Zurich, Switzerland.

Opach, T., & Nossum, A. (2011, July 3-8). *Evaluating the Usability of Cartographic Animations with Eye Movement Analysis.* Presented at the 25th International Cartographic Conference, Paris, France.

Paivio, A. (1986). *Mental representation: A dual coding approach.* New York, NY: Oxford.

Pashler, H. (1988). Familiarity and the detection of change in visual displays. *Perception & Psychophysics, 44,* 369-378.

Plaisant, C., Carr, D., & Shneiderman, B. (1995). Image-browser taxonomy and guidelines for designers. *IEEE Software, 12*(2), 21-32.

Plass, J. L., Homer, B. D., & Hayward, E. O. (2009). Design factors for educationally effective animations and simulations. *Journal of Computing in Higher Education, 21*(1), 31-61.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*, 179-197.

Rieber, L. P. (1990). Animation in computer based instruction. *Educational Technology Research and Development, 38*(1), 77-86.

Rieber, L. P. (2000). *Computers, Graphics, and Learning*. Athens, GA: Author.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review, 1*, 47-69.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events. *Perception, 28,* 1059-1074.

Spanjers, I. A. E., Wouters, P., van Gog, T., & van Merriënboer, J. J. G. (2011, January). An expertise reversal effect of segmentation in learning from animated worked-out examples. *Computers in Human Behavior, 27*(1), 46–52.

Swaak, J., & De Jong, T. (2001). Learner vs. system control in using online support for simulation based discovery learning. *Learning Environment Research, 4*, 217-241.

Tabbers, H. K., Martens, R. L., & van Merrienboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology, 74,* 71-81.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human–Computer Studies, 57,* 247-262.

Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology Review, 14*(3), 261-312.

Velez, M., Silver, D., & Tremaine, M. (2005). Understanding visualization through spatial ability differences. In *Proceedings of the IEEE Visualization 2005 Conference* (pp. 511-518). New York, NY: IEEE.

White, B. (1984). Designing computer games to help physics students understanding Newton's laws of motion. *Cognition and Instruction, 1*(1), 69-108.

Future Research and Development: Mapping the Terrain

V. FUTURE RESEARCH AND DEVELOPMENT: MAPPING THE TERRAIN

V. FUTURE RESEARCH AND DEVELOPMENT: MAPPING THE TERRAIN

Students' use of large scientific databases holds promise and may indeed be transformative for science education—but this territory is largely uncharted. New cycles of research and development are essential in order to create a map of "what works" (and what doesn't), why, how, and under what circumstances. We pose the research questions below (some of which emerged from our work on this KSR) as some overarching directions to pursue, recognizing that this list is far from exhaustive. These questions also can't be neatly parsed into separate endeavors: We can't evaluate how work with professional data sets improves student learning until we have well-designed interfaces and visualization tools that provide unfettered access to the data; conversely, the effectiveness of such interfaces can't be assessed unless they are co-developed with appropriate curricula and teacher supports. It will take the work of many researchers, curriculum developers and software developers along with iterative phases of development, design research, user testing, and pilot testing to realize the promise of professional scientific databases to transform science education.

The following questions are presented to map the terrain of research and development that is needed and to focus attention on certain areas that we believe will be particularly fruitful.

Authentic Data and Student Learning

Does work with authentic data sets (in the context of appropriately designed curriculum and teacher supports) improve student learning, and if so, how?

- What kinds of additional or characteristic learning occurs when students use professionally collected data sets, as opposed to concept-based visualizations or student-collected data?
- What kinds of additional or unique learning occur when students use real-time data sets that are accessible online?
- What is the impact on students' content-knowledge acquisition in science?
- What scientific practices or habits of mind are enabled for the student?
- What is the impact on students' ability to work with data visualizations, such as maps, graphs, and animations?
- What is the impact on students' development of cognitive skills, such as spatial reasoning?

Interfaces and Data Visualization Tools

What models of interfaces and data visualization tools (in the context of appropriately designed curriculum and teacher supports) would allow students to work most effectively with professionally collected data sets?

- What kinds of interface and visualization designs lead to increased engagement in germane cognitive load?
- How can we customize interfaces to individual users and then support these students as they grow in sophistication?
- How should an interface or visualization be adapted to suit particular types of data or data from a particular domain?
- How should an interface or visualization be adapted for particular types of tasks or learning goals?

- How do we overcome the particular challenges that students pose when creating their own data visualizations (e.g., unstructured investigations; messy, real-time data)?
- What kinds of talents, abilities, and tendencies do today's and tomorrow's students, who are tech-savvy in ways never before experienced, bring to the table?
- What tools are effective in developing students' spatial ability and helping them visualize 3D data (in particular, relating cross-sections to map views of 3D data)?
- How can interface designers find the right balance between overloading students with too many choices and providing them with sufficient choices to accommodate their particular needs?
- At what point do education users need to be involved in cyberinfrastructure development to ensure that the architecture supports the design of effective education interfaces? What kinds of conversations need to happen among educators, scientists, and technology developers?

Curriculum and Teacher Supports

While the Oceans of Data project has focused on the specifics of data interface design, this is but a small part of the ecology of effort needed to raise the scientific data literacy of the next generation of students. "Mining" large professional databases constitutes a new paradigm, not only for science research but for classroom instruction as well. None of the major science texts or curricula used in today's high schools involve students in accessing and working with large online databases, and activities requiring reasoning with data are rarely emphasized. Incorporating a new paradigm into science classrooms is a gradual process that require some adjustments. Investigating the ground truths of the science classroom, and supporting necessary change, is pivotal.

Here are a few important areas of focus that we believe are necessary in order to bring about transformative change:

- What kind of K-16 learning progression will lead to data-savvy adults and the scientific workforce of tomorrow?
- What models of curriculum and teacher facilitation help students transition from small, student-collected to large, professionally collected data sets?
- How can assessments be designed to measure students' learning as they build data literacy in science?
- What challenges does the integration of interfaces and associated curricula pose for pre-college teachers, who are likely to have limited experience in scientific research?
- What kinds of teacher training, professional development, and supports are effective in preparing pre-college teachers to facilitate students' work with professionally collected data sets?

Designing and refining interfaces and related curriculum, examining the impact on student learning, ascertaining teaching and classroom implementation factors, exploring the ways that usable data bases can affect what is taught and how, and collaborating with science database developers—these are all are interrelated components of the research and development agenda. The agenda underscores the need for comprehensive interdisciplinary efforts to chart the course for realizing the powerful educational benefits afforded by large scientific databases.

FOR MORE INFORMATION ABOUT VISUALIZING OCEANS OF DATA, **CONTACT:**

RUTH KRUMHANSL (rkrumhansl@edc.org), as co-director of EDC's Oceans of Data Institute, is engaged in efforts to bring authentic scientific data into the K–16 classroom. Krumhansl leads the Ocean Tracks project, a collaboration with Stanford, which provides student access to near real-time and archival data from electronically tagged marine animals, drifting buoys, and Earth-orbiting satellites, along with Web-based data visualization and analysis tools. Krumhansl played a lead role on Other Worlds/Other Earths, a project of the Harvard Smithsonian Center for Astrophysics enabling pre-college students to use online telescopes to detect planets orbiting other stars. In 2013, Krumhansl co-convened the EarthCube Education End-User Workshop, which developed recommendations for the National Science Foundation (NSF) regarding how to make digital earth science data broadly accessible to students and teachers.

Krumhansl is lead author of EDC Earth Science, a full-year high school earth science course (published in 2013 by LAB-AIDs). EDC Earth Science engages students in building from the evidence via data-rich investigations of questions that are relevant to society.

Before joining EDC, Krumhansl was a high school science teacher and department coordinator, a chief scientist in environmental consulting, and a petroleum exploration geologist. Her career in applied science immersed her in the search for patterns in complex geospatial data, providing a foundation for her current interest and work in preparing students to live in a data-intensive world.

CHERYL PEACH (cpeach@ ucsd.edu) is the director of Scripps Educational Alliances, a position focused on supporting the interplay between science and K–16 education at Scripps Institution of Oceanography. She is the Scripps PI for the Center for Ocean Science Education Excellence—California, co-PI on the Scripps GK–12 Scripps Classroom Connection program, education manager for the OOI Cyberinfrastructure Implementing Organization, and a program scientist at the Birch Aquarium at Scripps. Peach served a three-year term as the Education and Diversity subcommittee chair on the NSF Geoscience Directorate advisory committee and co-authored Geoscience Education and Diversity: Vision for the Future and Strategies for Success. Report of the 2nd Geoscience Education Working Group.

Peach hosted and co-convened the EarthCube Education End-User Workshop, and helped assemble the workshop report, providing the NSF and the EarthCube community with guidance on creating cyberinfrastructure that facilitates ready access to earth science data and data tools for use in education.

Prior to her tenure at Scripps, Cheryl was an oceanography faculty member and interim dean at Sea Education Association (SEA), where she instructed undergraduates in oceanography and oceanographic research, on shore in the classroom and at sea as chief scientist, on six-week research expeditions. At SEA, Peach was PI for Research at SEA, a five-year NSF professional development program that provided science teachers with a seagoing research experience.

EDC Learning transforms lives.

SCRIPPS INSTITUTION OF OCEANOGRAPHY
UC San Diego

This material is based upon work supported by the National Science Foundation under Grant Nos. 1020002 and 1019644. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.