Evaluating Digital Libraries for Teaching and Learning in Undergraduate Education: A Case Study of the Alexandria Digital Earth ProtoType (ADEPT)

Christine L. Borgman, Anne J. Gilliland-Swetland, Gregory H. Leazer, Richard Mayer, David Gwynn, Rich Gazan, Patricia Mautone

Abstract

This is a discussion on the research design for an educational evaluation of the Alexandria Digital Earth ProtoType (ADEPT), a digital library of geo-referenced information resources. ADEPT is being studied in undergraduate classrooms at the University of California, Los Angeles, and the University of California, Santa Barbara. The article provides a brief review of the deployment of digital libraries in educational settings, the role of information technology in developing students' scientific thinking, and the evaluation of digital libraries. We outline the overall research design, report on progress to date, and describe plans for the remainder of the five-year project. The article concludes with initial observations about classroom environments for using ADEPT and about the initial deployment of ADEPT prototypes.

Introduction

Digital libraries offer a wealth of opportunities to improve access to information resources in support of both "traditional" on-campus instruction and distance-independent learning (Borgman, in press). We are still at the early stages of realizing the potential of digital libraries in educational

Christine L. Borgman, Anne J. Gilliland-Swetland, Gregory H. Leazer, and Rich Gazan, Department of Information Studies, GSE & IS Bldg, University of California, Los Angeles, CA 90095-1520
Richard Mayer and Patricia Mautone, Department of Psychology, University of California, Santa Barbara, CA 93106
David Gwynn, Department of Geography, Bunche Hall, University of California, Los Angeles, CA 90095
Library Trends, Vol. 49, No. 2, Fall 2000, pp. 228-250

contexts, however. Few of the technological, logistical, and economic aspects of integrating digital libraries into university education have yet been assessed, much less the curricular and pedagogical challenges (National Research Council, Center for Science, Mathematics, and Engineering Education, 1998). Relatively little work has been done on evaluating the usability of digital libraries in any context, and minimal work has been done on assessing learning outcomes associated with the implementation of digital libraries in instruction. Many complex research design questions remain to be addressed, such as what to evaluate, by what methods, and how to determine if learning is occurring.

We report here on the research questions, research design, and preliminary observations from the first year of a five-year project (1999-2004) to develop and deploy a digital library of geo-referenced information resources ("geolibrary") in undergraduate courses at the University of California, Los Angeles, and the University of California, Santa Barbara. This study is part of the Alexandria Digital Earth ProtoType (ADEPT) project funded by the U.S. Digital Libraries Initiative, Phase 2 (National Science Foundation, 1999). ADEPT is an emerging digital library that will provide instructors and students with the means to discover, manipulate, and display dynamic geographical processes. The ADEPT system provides an interesting case study to observe the deployment of a digital library in instructional settings. Our thesis is that digital library services will contribute positively to undergraduate instruction and to student learning of scientific processes. To examine this thesis, we employ a variety of qualitative and quantitative methods to investigate the impact of ADEPT in undergraduate instruction. This article extends our initial reports on the education and evaluation component of ADEPT (Leazer, Gilliland-Swetland, & Borgman, 2000; Leazer, Gilliland-Swetland, Borgman, & Mayer, in press). Continuing reports will be provided on the ADEPT Web sites at UCLA (http://dlis.gseis.ucla.edu/adept/) and UCSB (http://www.alexandria.ucsb.edu/adept/).

Digital Libraries and Undergraduate Education

Educational applications of digital libraries range from primary school through graduate school and across all disciplines. One of our chief interests is how the use of digital libraries can promote thinking processes associated with problem domains (e.g., science, social sciences, humanities) at the undergraduate level. The first stage of the ADEPT educational evaluation focuses on scientific thinking and is being conducted in physical geography courses. This section provides a brief literature review of the role of digital libraries in education and in scientific thinking to set the context for the case study of ADEPT. The literature is reviewed relating to the evaluation of digital libraries in general.

Educational Applications of Digital Libraries

Faculty and librarians alike are concerned about ways to implement digital libraries in education. The Council on Library and Information Resources (1999) held a meeting "to consider changes in the process of scholarship and instruction that will result from the use of digital technology and to make recommendations to ensure that libraries continue to serve the research needs of scholars." Among their recommendations was that institutions of higher education should "place more emphasis on training and support for faculty use of information and instructional technologies."

The University of Michigan Digital Library Project (Wallace, Krajcik, & Soloway, 1996) posits that the main benefit of digital libraries in the classroom is improved means and opportunity for inquiry-based learning. A component of this research, the Middle Years Digital Library project (Soloway et al., 2000), allowed science students in grades six through nine to learn and explore topics in a less-regimented manner than traditional textbook learning. The few outcome-based studies that have been conducted have suggested a positive correlation between integrating electronic information sources into the classroom and increased scholastic success. Newnham, Mather, Grattan, Holmes, and Gardner (1998), for example, gave geography students access to Internet source material downloaded onto a local network file server. Students were encouraged to make use of the material and communicate among themselves via electronic mail. The study found that access to electronic geo-information sources enhanced student learning.

In a distance-learning study, Mose and Maney (1993) found that geology students with access to a combination of televised instruction and computer-based communications software demonstrated higher levels of learning than did non-computer-equipped students. Data were collected via software questionnaires, case study interviews, and course grades. While this study focused on the student-student and student-instructor communications opportunities provided by educational technology, the researchers concluded that the use of this technology in undergraduate geology courses facilitated learning and increased student engagement.

At the other end of the age spectrum, digital libraries also have helped very young children understand complex scientific concepts. Many complex concepts become understandable when taught in a contextualized and incremental manner (Metz, 1995). Kafai and Gilliland-Swetland (in press) built on Metz's work in a study where young science students recreated the process of generating and describing digital scientific documentation by emulating the activities of an early naturalist. The researchers noted that the students found visual materials more intellectually accessible, even when the source materials were meant for more advanced students.

Digital Libraries and Scientific Thinking

One method of determining the success of digital libraries in improving student learning is to examine whether they are helping to achieve pedagogical objectives. One such overarching objective in geography instruction, for example, is the development of scientific thinking in students. Consensus exists that students need to learn five skill sets in order to engage in scientific thinking in geography: (1) asking geographic questions, (2) acquiring geographic information, (3) organizing geographic information, (4) analyzing geographic information, and (5) answering geographic questions (Geography Education Standards Project, 1994, pp. 42-44).

The first skill set—asking geographic questions—involves being able to pose questions that can be addressed in the field of geography. When faced with an issue, students need to be able to formulate geographic questions, such as How did that get there? or What are the consequences of that being there? Digital libraries have a role to play at the question-asking phase of scientific thinking, because students can get ideas for questions by browsing some of the available information. In this way, the browsing capabilities of digital libraries can aid the user's question-asking process.

The second skill set—acquiring geographic information—includes locating and collecting relevant information, such as reading maps and other visual representations of space. The information-gathering phase represents perhaps the most obvious venue for digital libraries, because they can greatly increase the efficiency of obtaining relevant information and may even allow access to information that would not otherwise be available. In this way, the information retrieval capabilities of digital libraries can aid the user's information-gathering process.

The third skill set—organizing geographic information—includes systematically arranging and displaying geographic information, such as maps and graphs. For example, in some cases, a user may create a map based on collected information. Digital libraries can help users in their information-organizing process if the libraries include tools for manipulating information—especially tools for creating visual representations of information.

The fourth skill set is analyzing geographic information—a process that includes finding patterns, trends, relationships, and connections in the information one has gathered and organized. In some cases, the analysis may involve scrutinizing patterns in maps or charts, and in other cases the analysis may involve statistical analyses of quantitative data. Digital libraries can aid in this information-analysis process by providing tools for aggregating and analyzing data.

The fifth skill set is answering geographic questions and often takes the form of a written or oral generalization or conclusion. The *Standards* emphasize that "students should also understand that there are alternative ways to reach generalizations and conclusions" (Geography Education

Standards Project, 1994, p. 44). Digital libraries can assist in this question-answering process by allowing users easily to check the predictions of the explanatory models they construct.

By understanding how students will use a digital library in the context of scientific thinking, it is possible to construct the digital library in a way that will support the underlying processes. In short, digital libraries are more than storehouses of information; they should be aids to the question-asking, information-gathering, information-organizing, information-analyzing, and question-answering processes of users. In this way, digital libraries can also support the broader call in the *National Science Education Standards* (National Research Council, 1996) for allowing "inquiry into authentic questions generated from student experiences" (p. 31). Although we focus on promoting geographic thinking in our project, the same kinds of skills that support geographic thinking also apply to other scientific disciplines. The five skills can be used in inductive (or data-driven) reasoning, such as looking for trends in data that lead to a theory, or in deductive (or theory-driven) reasoning, such as testing two competing theories through dynamic modeling.

Evaluating Digital Libraries

The goals of the ADEPT project, from an educational perspective, are to construct a digital library that will make geo-spatial and geo-referenced information resources useful in undergraduate instruction and whose use will lead to better learning outcomes than with traditional modes of instruction. We are conducting both formative evaluation that assists in formulating design requirements and summative evaluation that assesses learning outcomes from using the system in instructional settings.

Digital libraries are difficult to evaluate due to their richness, complexity, and variety of uses and users. Few proven methods are available. The need for evaluation methods and metrics was among the key findings of the *Social Aspects of Digital Libraries Workshop* (Borgman et al., 1996). Some progress is being made, as evidenced by this special issue and by a forthcoming book on the evaluation of digital libraries (Bishop, Buttenfield, & Van House, in press). Most evaluation studies of digital libraries address questions of usability (Borgman, in press-b). "Usability," however, like "user friendly," is an amorphous term with a wide range of context-dependent interpretations.

Many general criteria for usability exist, such as those proposed for "every citizen interfaces to the nation's information infrastructure" (Toward an Every-Citizen Interface..., 1997, p. 45). Criteria include "easy to understand, easy to learn, error tolerant, flexible and adaptable, appropriate and effective for the task, powerful and efficient, inexpensive, portable, compatible, intelligent, supportive of social and group interactions, trustworthy—secure, private, safe, and reliable, information centered,

[and] pleasant to use" (p. 45). Other applicable criteria are the user interface design rules established by Shneiderman (1998) as adapted to information retrieval (Shneiderman, Byrd, & Croft, 1997): strive for consistency, provide shortcuts for skilled users, offer informative feedback, design for closure, simple error handling, permit easy reversal of actions, support user control, and reduce short-term memory load. Nielsen (1993) identifies five general usability attributes for information systems as well as other applications: learnability, efficiency, memorability, errors, and satisfaction.

These principles offer general guidance for design but are far from a "cookbook" for constructing any individual digital library. Principles such as "easy to learn" must be applied relative to the application and the user community. A system that supplies daily weather reports to the public must be much easier to learn than one that supplies geophysical data to researchers, for example. Setting appropriate benchmarks for any given system involves evaluation with members of the target audience and comparisons to similar applications.

Design guidelines and evaluation criteria can be employed to build more usable systems but only to the extent that design goals are appropriate for the application. Determining appropriate design goals for digital libraries is itself a challenge given the early stages of research on uses, users, and usability and the rapid evolution of the underlying technologies. Formative evaluation is a particularly valuable approach in such situations, because user needs and requirements can be studied concurrently with initial stages of designing the system (Gilliland-Swetland, 1998; Marchionini & Crane, 1994; Fitz-Gibbon & Morris, 1987).

The Digital Library for Earth Sciences Education (DLESE) project (http://www.dlese.org/SoftwareArchitecture/requirements/index.html), with which ADEPT is a cooperating partner, also is in the early stages of formative evaluation. DLESE has solicited scenarios from their user community of K-12 teachers as a basis for identifying requirements for functionalities and systems architecture. The scenarios are reviewed and refined with assistance from the user community, which is actively involved in the project. The DLESE project is finding that the requirements are a moving target because, as community level of sophistication grows, so does the number and sophistication of their requests. One goal of the evaluation plan is to examine how well the final products meet the functionalities defined in the scenarios submitted by the community (Marlino, personal communication, August 14, 2000).

ADEPT Undergraduate Implementation and Evaluation

ADEPT is an extension and enhancement of the Alexandria Digital Library (ADL), which was developed at the University of California, Santa

Barbara (UCSB) under the first Digital Libraries Initiative (1994-1998). ADL is an operational digital library that provides access to collections of maps, images, and other geo-referenced materials from a 1.5 terabyte (and growing) collection of materials from UCSB's Map and Imagery Laboratory (http://www.alexandria.ucsb.edu). The operational version of ADL provides users with access to services that allow them to answer such questions as what information is available about a given phenomenon at a particular set of places. ADL also provides new types of library services based on gazetteers and other information access tools. ADL went online in Fall 1999 as part of the California Digital Library (http://www.cdlib.edu). Formative evaluation of ADL focused on multiple target user communities (earth scientists, information specialists, and educators) using a variety of methods, including online surveys, ethnographic studies, a classroom study with a later version of the interface (Hill et al., 2000) and transaction logs (Buttenfield & Kumler, 1996).

The ADEPT project, also centered at UCSB, is developing a digital earth metaphor for organizing, using, and presenting information at all levels of spatial and temporal resolution. A central aspect of ADEPT is the development of I-scapes (information landscapes), whose working definition is as follows:

> I-scapes are a means of expressing and visualizing geo-spatial concepts and processes, for research, instruction, and learning. I-scapes also include a set of tools and resources for the use of geo-spatial and geo-referenced information resources in teaching undergraduate courses. Instructors will employ I-scapes to convey concepts; students will use I-scapes to learn concepts and to get experience manipulating information resources in the ways that domain experts use them.

Our working scenario for the use of ADEPT I-scapes is that the course instructor will define the scope and concepts of a topic to be taught. The instructor, with the aid of a graduate student researcher from the education and evaluation team, will assemble a small collection of information resources for teaching the topic and will apply ADEPT tools and services to create I-scapes. The instructor will use one or more I-scapes to present the topic in class lecture sessions. Teaching assistants also will use I-scapes to discuss and demonstrate the topic in laboratory sessions. Students will perform exercises in lab sessions and outside of class using I-scapes to test hypotheses in the pre-selected collection of resources.

In support of this proposed scenario, the ADEPT project is developing a range of analysis tools and modeling services that will enable users to construct their own personalized digital libraries and to use them in creative ways alone and in collaboration with other users.

At the core of effective digital library design is the relationship between the content to be provided and the user community to be served. Design goals can originate from either perspective. Best of all, design goals from both perspectives should converge. We are taking a convergence approach to design, with the education and evaluation team focusing on needs assessment, evaluating prototypes in active use, and identifying system requirements. Concurrently, the ADEPT implementation team is focusing on evolving the ADL testbed architecture and services, such as interface specifications, service prototypes, interoperability, and collection growth and diversity. Ours is an iterative and collaborative approach to development, with evaluation integrally embedded in design. Needs are identified from the user and collections perspective, prototypes are constructed and evaluated, and the results fed back into the design and development process.

Research Questions

We hypothesize that digital library services that provide instructors and students with the means to discover, manipulate, and display dynamic geographical processes will contribute positively to undergraduate instruction and to the development of scientific and other discipline-specific reasoning skills. This hypothesis generates a number of research questions, only a few of which are addressed in this article. Here we focus on how the evaluation research design addresses the following questions:

- How can ADEPT modules support domain knowledge, work practices, and reasoning models of multiple disciplines that use geo-spatial resources? For example, can ADEPT modules and services be structured in such a way that they will help a student to think like a geographer or an environmentalist?
- How can ADEPT accommodate users with different skills, knowledge, cognitive styles, and pedagogical styles? While it is difficult for any digital library to support such heterogeneity in its users, are there ways in which ADEPT can facilitate moving between different domain knowledge and technological skill levels as users become more sophisticated? Are there also ways in which users who are less comfortable with the spatial metaphor can enhance their spatial processing capabilities?
- How can ADEPT help users view primary geographical evidence in new ways to answer scientific or geographical questions? ADEPT provides users with diverse primary research data (e.g., remotely sensed data) as well as published information. It also provides users with links to non-digital information held elsewhere and enables users to incorporate additional content that they have created or collected themselves. Using the tools and services provided by ADEPT, can users combine, manipulate, and visualize these resources in ways that will allow them to ask and answer questions in original and creative ways?
- How can ADEPT support the range of heterogeneous resources and their metadata necessary for learning applications? Even though ADEPT's holdings are vast, it is difficult to anticipate the specific

resources that instructors or students may wish to incorporate into a class presentation or project. Frequently, they will wish to draw upon additional materials that they already have in their possession, materials that may not be in the preferred file formats or accompanied by the systematic metadata of existing ADEPT holdings.

Project Methods

We are addressing these research questions through several concurrent approaches, including establishing general design principles; analyzing cognitive processes and information-seeking needs of instructors and learners; analyzing practices, behaviors, and knowledge-processing requirements of the disciplines within which ADEPT is being implemented; and developing case-based prototypes.

*General Design Principles.*We began the project with a top-down approach to setting requirements, establishing general design principles such as:

- ADEPT should support real scientific problems that can be studied, learned, or solved with the use of geo-spatial and geo-referenced information resources.
- I-scapes should focus on dynamic (rather than static) processes for which "real data" can be visualized and manipulated by instructors and students.
- Research should concentrate initially on geography, as we have the most knowledge and information resources in this domain.
- Research should expand later to other disciplines that use geo-spatial and geo-referenced information resources (e.g., geology, earth and environmental sciences, sociology, urban planning, and even humanities fields that sometimes organize content by geographic location, such as art history and theology).
- ADEPT must be easy to learn and use by undergraduates with minimal domain knowledge and technical skills (e.g., freshmen in geography courses for non-majors).
- ADEPT must be easy for instructors to learn and use.
- ADEPT should improve teaching productivity.
- ADEPT should create minimal additional workload for instructors.

The latter three principles address the problem of incentives for faculty to use ADEPT in instruction. While faculty at research universities such as UCLA and UCSB are held to high standards of teaching, it is but one of several significant demands on their time. They are most likely to adopt new teaching tools and methods if the overhead in time and effort is minimized and if the resulting advantages are deemed worthy of the investment. The project includes a substantial amount of graduate student research assistance for working with faculty in developing and deploying I-scapes in instruction as a means to encourage participation.

User Groups and Tasks. Based on the usage scenario described earlier, we identified four user groups for study: (1) faculty (in their role as course instructors), (2) teaching assistants, (3) students in the courses where ADEPT is implemented, and (4) students continuing to study the discipline after the completion of the course. These groups differ in needs, activities, and levels of domain knowledge. Analysis of baseline data indicates that each of these user groups generates different requirements for what and how ADEPT resources and services are developed. We have identified the following list of candidate tasks that ADEPT should support:

- highly-directed uses such as lab exercises to reinforce a specific disciplinary concept;
- instructional modules that introduce concepts in an incremental manner and can be customized and extended by faculty for use in lectures;
- free-form exploration conducted by students preparing term papers or faculty putting together a lecture that might include personal manipulation of data sets, information visualization, and the integration of new information or data sets to augment existing content;
- collaborative applications that might be used by students doing team projects or faculty and teaching assistants who are team-teaching; and
- discipline or domain-specific methods of building knowledge that support specific information seeking and use processes.

Information systems and services designed to facilitate learning must accommodate a variety of pedagogical goals and styles and a variety of learning styles. The *processes* whereby information is identified, selected, retrieved, manipulated, annotated, and presented to others often are more important than the retrieval and use of the information itself. To examine these usage contexts and the interplay among them, we are conducting three types of studies, employing qualitative and quantitative methods: classroom-based studies, laboratory studies, and system use studies. Some will be conducted longitudinally and others will occur at different points throughout the development of ADEPT.

Assessing Learning Outcomes. How can we assess students' scientific thinking in geography? What cognitive changes occur in students who receive experience in ADEPT that do not occur in non-ADEPT students? In short, what are the cognitive consequences of participating in an ADEPT environment? Our evaluation of cognitive learning outcomes seeks to answer these questions through performance assessments of scientific problem solving. Performance assessment involves giving learners realistic tasks and carefully observing how they go about handling them; in a science domain, this means presenting a scientific problem and observing how students engage in scientific problem solving—that is, observing how students actually "do science" (Doran, Lawrenz, & Helgeson, 1994, p. 415). Performance assessment techniques are popular in science education because they allow for a richer assessment that goes beyond testing for students' remembering specific facts (Baxter, Shavelson, Goldman, & Pine, 1992; Persky et al., 1996). We plan to assess the cognitive consequences of participation in the ADEPT program by testing an ADEPT group and a comparison group on a series of performance tasks, each tapping one of the five target skills in geographical thinking.

1. *Asking Geographic Questions.* The first skill is to formulate a testable question. To assess this skill, we will present students with a geography scenario, expressed as a short video, and ask them to generate as many testable questions as possible. Scoring will be based on the number of acceptable questions that each student proposes.
2. *Acquiring Geographic Information.* The second skill is to gather relevant information. To assess this skill, we will present students with a geography question, expressed as a short video, and ask them to list the kinds of information they would need. As in the prior test, scoring will be based on the number of acceptable information requests that each student proposes.
3. *Organizing Geographic Information.* The third skill is to organize relevant information in a way that supports scientific thinking. To assess this skill, we will present various information sources to students in the form of text, video, or graphics files that can be accessed and ask them to create summary graphics for a future presentation about a target question. For this and the following two tests, scoring will be based on a scale of 0 to 5.
4. *Analyzing Geographic Information.* The fourth skill is to find patterns or relations in organized geographical material, such as graphics. To assess this skill, we will present various summary graphics—intended to address a geography question—and ask students to write a sentence to accompany each one.
5. *Answering Geographic Questions.* The final skill is to create a verbal conclusion or generalization to a target question based on organized and analyzed information. To assess this skill, students will be asked to write a one-paragraph answer to a target question based on a series of narrated graphics that organize and analyze the relevant information.

The design of a performance assessment program involves a number of issues. Our design decisions are guided by a conception of assessment in which the quality of a test depends on four characteristics:

1. *reliability*—the test gives a consistent score,
2. *validity*—the test measures what it is supposed to measure,

3. *standardization*—the test score allows for comparison among test-takers, and
4. *objectivity*—the test is scored and administered the same way for everyone.

First, should we focus intensively on one geography scenario or broadly on several? We plan to include several geography scenarios rather than focus on one because a broader set of scenarios is likely to increase the reliability of our assessment.

Second, should we test for near transfer (i.e., problems like those used during instruction) or far transfer (i.e., problems that are not closely related to those used during instruction)? We opt for testing students on problems that are similar in format to the problems used during instruction so they require the same scientific thinking skills but which involve different geography content so students cannot simply remember specific answers. This approach increases the validity of our assessment.

Third, should our measurements be quantitative—i.e., in the form of numbers—or qualitative—i.e., in the form of a written summary of our observations? We opt for quantitative measurements that are tied to a clear scoring rubric. This approach increases the standardization of the assessment.

Fourth, should we provide scripted or open-ended guidance to students as they seek to solve the problems we present? We intend to provide scripted guidance so that all students will receive the same kinds of interactions with teachers. This approach increases the objectivity of the assessment.

The final product will be a set of performance assessment instruments that allow for quantitative measurement of each of the five target skills in scientific thinking in geography.

Research Schedule and Strategy

First Year Progress (1999-2000). This first academic year of the implementation and evaluation component of the ADEPT project has been devoted to requirements analysis, evaluation design, and pilot testing that is concurrent with the development of the ADEPT architecture by the UCSB-based development team. Activities have included:

- establishing general design principles;
- developing the evaluation design and instruments;
- identifying and recruiting faculty and students to participate in implementation and follow-up;
- identifying pedagogical goals and styles in participating faculty through classroom observations and interviews;

- identifying canonical concepts (from texts, standards, interviews with faculty, and examination of instructions to teaching assistants) that can be introduced incrementally and described by flexible instructional modules and used at different levels of granularity;
- developing prototype I-scapes and pilot testing our evaluation methods in courses at UCLA and UCSB;
- gathering baseline demographic and performance data on students who have taken the classes under study in the previous five years as well as students currently enrolled in the classes; and
- gathering preliminary feedback from students at the mid- and endpoints of the classes in which prototype modules of ADEPT have been implemented.

Based on consultations with the chairs of the geography departments at UCSB and UCLA, we determined that the introductory courses in physical geography at both campuses were best suited for initial studies. This course is taught three times per year at UCLA (Fall, Winter, and Spring terms), which is a larger campus, and the equivalent course is taught once per year at UCSB. The course is taught at the lower division level (freshman-sophomore) and satisfies general education requirements for the bachelor's degree, so it draws students from all disciplines. The course also is a prerequisite for geography majors. All four instructors agreed to participate in the evaluation study, and all four sections of the course were observed on a regular basis by members of the ADEPT education and evaluation team. The instructors also provided copies of laboratory assignments and exams used to assess students.

During the Spring term (April to June 2000), prototype I-scapes were deployed in the physical geography courses at UCSB and UCLA. We consulted the instructors to identify a list of course topics that involved dynamic processes and that could be explained better through dynamic presentations rather than the current static presentations on overheads, slides, or chalkboards. Of the suggested list, the topics selected for the initial I-scapes were hydrology and fluvial processes, as a body of materials and instructor expertise were most readily available. At UCSB, four ADEPT lectures were guest lectures given by a faculty member who is part of the ADEPT team rather than being presented by the course instructor.

Plans for Year Two (2000-2001). Based on knowledge gained from the first year of the project, we will devote the second year to a case-based bottom-up approach to the education and evaluation project. We are constructing several exemplar cases that will help to generalize design from specific instances. We have selected two cases from those previously agreed upon by participating instructors of physical geography and one for a course on human geography to be taught in the 2000-2001 academic year. The physical geography cases are erosion, as a topic on fluvial processes, and subduction, as a topic on plate tectonics. For human geography, we have tentatively selected Von Thünen models as a topic on land usage.

We will collect appropriate content for these topics, working with participating instructors, and the implementation team will develop services and functional capabilities for ADEPT prototypes. To date, we have identified these requirements for the three cases:

- appropriate metadata and representation of content;
- appropriate searching capabilities to select content within I-scapes;
- ability to manipulate appropriate parameters to demonstrate processes and test hypotheses;
- visualization features to demonstrate processes and test hypotheses;
- ability for individual students and instructors to save their work for reuse; and
- instrumentation to capture user-system interactions.

Evaluation will be conducted in classrooms and in laboratory settings, as discussed earlier.

Plans for Years Three through Five (2001-2004). Years three through five of the project will continue the usability and evaluation studies with subsequent iterations of I-scapes in multiple classrooms in multiple disciplines. Lectures are only a starting point, as noted earlier. We plan to expand ADEPT into laboratory applications, independent learning, and provide functionality for collaborative and distributed learning.

Initial Observations

Our findings are preliminary, as we are still analyzing the first year's data at this writing, and ADEPT itself is in the early stages of development. We report here only on baseline classroom observations and on results from the prototype I-scapes deployment at UCLA and UCSB. The prototypes provided a first look at introducing computer-based technology into geography classrooms for teaching dynamic processes and an opportunity to test our instruments. We have no learning outcome data yet. Further data from faculty interviews, student interviews, and student demographics will be reported later. Initial observations reported here are in two categories: classroom environment and implementation of ADEPT I-scapes prototypes.

Classroom Environment

We recognized early in the project that ADEPT modules would have to be flexible, adaptable, and relatively small in scope. Among the design questions we are exploring at this stage are the following: At what level of detail or granularity should I-scapes be created? How detailed can or should modules be if they are to be useful for multiple instructors? What are the implications for metadata to describe and represent the concepts? What are the collection requirements for ADEPT? These questions guided our observations of lectures and discussion sections of the physical geography courses at UCLA and UCSB. Faculty and graduate student researchers from the ADEPT education and evaluation team observed a sampling of class sessions before, during, and after the implementation of the I-scapes prototypes. We also videotaped one lecture by each of the UCLA instructors and videotaped the I-scapes lectures at UCSB. These observations form baseline data to determine what is in common and what varies across the multiple offerings of the same and similar courses, what aspects of the courses might be incorporated in I-scapes, and what aspects are independent of I-scapes. We are not attempting to reform undergraduate education as a whole; rather, our purpose is to identify ways in which digital libraries can be utilized effectively in instruction.

At UCLA, the same introductory physical geography course was taught by three different instructors in one academic year (Fall, Winter, and Spring terms on the quarter system), providing the opportunity to compare multiple approaches to teaching the same course. The comparable introductory course at UCSB, taught only once per year, has similar content and used one of the same textbooks as at UCLA. The three UCLA courses were taught with two lectures per week plus one required laboratory session taught by a graduate teaching assistant. The UCSB course was taught with three 50-minute lectures per week plus one required laboratory session taught by a graduate teaching assistant. Enrollment ranged from 60 to 120 students, and four to six laboratory sections were offered for each course. In total, we observed four courses (three at UCLA and one at UCSB) and five instructors (three at UCLA and two at UCSB: the regular course instructor and the guest who presented the four I-scapes lectures).

Course Content. Topic emphasis varied considerably due to differences in course texts and in instructors' interests and expertise. The five instructors were experts in climatology, geomorphology, remote sensing, river systems, pedology, and soil evolution. Although all the instructors covered all the requisite topics, the proportional amount of time devoted to each topic reflected their respective research areas. In interviews with the instructors and in observation of their lectures, all had a core teaching goal in common, which was to present geography as a system of interacting processes. In the areas of their greatest expertise, instructors were most likely to interject anecdotal stories, specific case studies, and state-of-the-art research into the classroom discussion. In doing so, instructors were able to personalize the scientific process and explain "how science is done" in the field or in the lab.

The three UCLA instructors chose three different textbooks and followed them to varying degrees. One instructor followed the text closely, lecturing chapter by chapter and testing students on textbook content.

Another drew illustrations from the text and incorporated other materials into lectures, testing students on all materials covered in lectures and lab sessions. The third instructor relied on the text for background reading and structured his lectures much differently than the textbook organization. The UCSB instructor and guest lecturer occasionally referenced textbook chapters for the day's lecture but did not follow the text closely, nor did they follow the ordering of the chapters. From time to time, one of the UCSB instructors would indicate parts of the textbook, such as a table or graph, which summarized something he had just discussed, or he would display a table pulled from the text and elaborate upon it.

Teaching Styles. We also found considerable variation in teaching style, part of which we attribute to the amount of teaching experience. At UCLA, one of the instructors was an assistant professor who was teaching the course for the first time, another was a recently-promoted associate professor who had taught the course at least once before, and the third was a senior full professor who had taught this course many times over a period of several decades. At UCSB, the instructor and guest instructor were full professors with many years of teaching experience, though neither had taught this course recently. The five faculty varied in lecture style, use of instructional technology, and their approaches to engaging students in discussion.

Perhaps, as might be expected, the degree of reliance on the textbook varied with teaching experience. The assistant professor relied most heavily on the text and the full professors the least. While all the instructors answered student questions during lectures, the degree to which they directly elicited student involvement varied. At UCLA, one instructor actively involved students in the lecture sessions, while the other two tended to defer student discussion to office hours and lab sections, devoting more classroom time to lecturing. At UCSB, the instructors would occasionally direct questions to the students or give them a scenario and ask them to make predictions, but they also tried to engage the students by drawing references between the topics being discussed and students' personal experiences. For example, in discussing types of land or river formations, the instructor asked how many students had taken a hike in a certain local area, asked them to describe what they saw, then related specific features of that area to a topic in the lecture. For the most part, however, any active student discussion took place in lab or during office hours.

The five faculty also had different approaches to the use of instructional technology in their classrooms. At UCLA, one instructor lectured almost entirely from notes on an overhead projector, another lectured almost entirely from the chalkboard, and the third used a mixture of chalkboard and overheads. The instructor who made the most use of the chalkboard wrote copious notes on the board with detailed diagrams. Dynamic processes were illustrated in multiple colors of chalk. He often came to the classroom about twenty minutes before class to prepare his diagrams on the board. He also brought in a large map of the world on which he would point out the location where specific processes occurred. Some lectures were augmented with slides and one with rock samples.

At UCSB, both instructors primarily used computer presentations which were displayed on one or both of the two screens in the front of the room and on several smaller television screens mounted throughout the lecture hall. The primary course instructor used a hypermedia format for his computer presentation, which usually consisted of an outline of key terms and an accompanying image or animation. He occasionally used an overhead projector, displayed on the second screen, to write down additional information or to use one of his favorite transparencies. He did not read from prepared notes but instead used the outline of his computer presentation as a prompting tool. He would spend from two to twenty minutes on each screen image. The guest lecturer used two computers, presenting his lecture outline on one screen and images and animations on the other. His lecture was fairly fast paced, mostly following the outline but occasionally deleting or re-ordering topics.

The instructors also varied in the emphasis they placed on learning specific concepts and on learning processes. One instructor focused on approximately six processes per lecture, each illustrated with an overhead from the textbook. This format allowed students to follow the lecture in the book and to repeat concepts and clarify processes during the lecture period.

We also noted differences in the instructors' use of the physical space in the classroom. One of the UCLA instructors was very active, constantly moving as he drew detailed diagrams at the board and as he pointed out locations on the world map. Another instructor was moderately active, as she switched from writing notes on the board, to speaking from the overhead, to entertaining questions from the class. The latter two instructors were in a large bright lecture hall with many windows and a high ceiling. The third instructor was least active, standing at the podium for the duration of the lecture, showing overhead figures and writing on transparencies. His course was held in a low-ceilinged windowless room that he kept dark to maximize the legibility of his overheads.

The two UCSB faculty were physically constrained by their use of computing equipment, although they used it differently. The regular professor used one computer terminal, which was located in the lectern. He would pace around the front of the lecture hall, point to relevant areas on the projection screen, and speak directly toward the students, returning to the lectern only when he needed to switch to the next screen. The guest professor occasionally walked up closer to the students and moved around but, due to the use of two computers, the faster pace of screen changes, and some technical difficulties with the microphone, was more restricted in physical movement. He wrote diagrams and notes on the computer screen using a computer pen rather than using an overhead projector.

Classroom Implementation of ADEPT I-scapes Prototypes

Given the amount of development required to produce fully-operational I-scapes with real scientific data, we took a rapid-prototyping approach and constructed simple prototypes using MicroSoft PowerPoint. This approach enabled us to incorporate text, photographs, diagrams, images, and moving images that illustrated dynamic processes and allowed the instructor to add comments and annotations in real time. As noted in the prior section, the UCLA instructor developed one lecture on fluvial processes using a one-screen method that combined graphics, simulations, and the lecture outline, while the UCSB guest instructor developed four lectures on hydrology and fluvial processes using a two-screen projection method (one for graphics and simulations and one for the lecture outline). Both instructors were assisted in the I-scapes development by graduate student researchers employed by the ADEPT project.

Following are some of our initial findings on the classroom implementation in Spring 2000 and their implications for subsequent iterations of ADEPT.

Integrating Information Resources into ADEPT I-scapes. Despite the richness of the Alexandria Digital Library, we found it necessary to locate additional information resources in support of the hydrology and fluvial processes lectures for introductory physical geography courses. Instructors wished to integrate materials such as lecture outlines and notes, diagrams, or personal slide collections for which they presumably held intellectual property rights. Other materials of interest were drawn from textbooks, CD-ROMs, online resources, or other sources for which they did not hold the rights. The need to integrate additional materials has implications for system design, management of intellectual property, and sharing of resources. Under fair use guidelines, instructors normally can present published materials in a classroom lecture and often do. When materials are incorporated in other products, posted online, or shared with other instructors, rights and permissions are much less clear. The simple prototypes were developed for research purposes only and will not be shared until and unless we can resolve the intellectual property issues.

Presentation Capabilities. In gathering materials for the ADEPT I-scapes prototypes, we found that display, layout, and other presentation features are essential considerations. Instructors often selected illustrations based on graphical qualities over relevance and familiarity (e.g., an image of a river in Africa was visually more striking than an available image of a local river).

Visual context must be provided by clear labeling, zooming, use of recognizable geographic features, and other means.

Instructors' Experiences with Information Technologies. All of the instructors studied at UCLA and UCSB are accomplished researchers who employ high-end technology in their scholarship. Geography is a technology-intensive field, particularly in areas covered by courses such as climatology and geomorphology. Even so, the UCLA instructors did not normally employ computer-based instructional technology in teaching this introductory course (other than that provided by the ADEPT project), instead relying upon chalkboards, overhead projectors, slide projectors, and paper maps for instruction. When asked their reasons, they said that too much advance planning was required for computer-based instruction, and that too much assistance would be required to install equipment, keep it running, and so on. They were interested in experimenting with new instructional methods, however, and were willing to participate in ADEPT because we supplied equipment, technical assistance, and graduate assistant support in developing lecture materials. Several of the instructors commented that they would prefer, at least initially, to have "canned" materials rather than live digital libraries or online connections in the classroom. Two faculty members commented that they did not wish to present a technology-based lecture in front of 100 or more students "without a net." The course section studied at UCSB did incorporate computer-based materials and had technical assistance at a level equal to, or greater than, that supplied by ADEPT.

The UCLA instructor who implemented the ADEPT I-scape prototype Spring term told us in interviews that he was willing to experiment with our technology in his lectures, acknowledging that he was somewhat apprehensive about the computer-driven nature of the presentation as it ran counter to his usual teaching style. He found the experience satisfactory, however, and sees ADEPT's primary benefit as an effective visual aid to communicate concepts in physical geography. While he was concerned initially that the extra effort and stress might be detrimental to his teaching, he felt that his students benefitted from the experience. He found the one-screen I-scapes module somewhat cumbersome, however. He had trouble moving back and forth through concepts as he does in regular lectures and felt constrained by the computer. He also felt the slide show sped up the lecture and did not give students time to digest the content and to take notes at the same time. He said that next time he would keep overheads nearby and use ADEPT as a supplement, not as the driver of a whole lecture.

The UCSB guest instructor used dual screens, one for the illustrations and one for the lecture outline. He was very positive about the experience (it "converted him"), even though he had not used computer-based tools in his own teaching in the past. He said that he "loves it" and "feels like he can get his points across more cleanly and effectively." He also thinks that the only way for students to comprehend complex phenomena is to see these demonstrated visually. He did find that ADEPT was something of an obstacle to communicating with students because the technology drew some of his focus, but thought that he would become more comfortable as he became more familiar with the setup. Also, the computer kept him physically in the same place during the lecture, which he sees as an improvement over the amount of pacing across the room he normally does during lectures.

Conclusion

Our goals for the ADEPT project are to construct a digital library that will make geo-spatial and geo-referenced information resources useful in undergraduate instruction, ultimately leading to better learning outcomes than with traditional modes of instruction. The first year of the education and evaluation component of the ADEPT project has been devoted to establishing general design principles; developing the research design; developing and pilot testing data collection instruments; gathering baseline data on how geography courses currently are taught and on the students who take these courses; developing, deploying, and evaluating the first ADEPT I-scapes prototypes; and planning subsequent stages of the project.

The research design, including the development of methods and instruments for gathering qualitative and quantitative data, has been a substantial undertaking in itself. To date, we have drawn our methods from educational evaluation, cognitive psychology, human factors, systems analysis, and user-centered design. We expect the methods and research questions to evolve throughout the project as the technology, classroom environments, and user requirements are all moving targets.

Our initial observations suggest that matching the content and capabilities of I-scapes to the range of instructors' approaches to teaching the same topics will be a considerable challenge. The five instructors we have studied so far vary substantially in their topic emphases, choice of texts, use of instructional technology, and student assessments. We need to determine the appropriate granularity of I-scapes topics as well as their contents and their features, such as the abilities to manipulate data, test hypotheses, and visualize processes. The early data also suggest that the presentation quality of images or simulations may be as important as their source or location. These instructors traded clarity and labeling of images with familiarity to the students (e.g., local rivers versus rivers on other continents). Similarly, the metadata to describe content is an issue. While some information resources were sought by location (e.g., names of local rivers; latitude and longitude coordinates), others were sought by type and topic (e.g., simulations of river erosion). Pacing is yet another concern.

Much more material can be presented with computer-based instructional tools than with chalkboards and overhead projectors, and students easily can be overwhelmed. Conversely, students may be entertained by slick presentations without learning the scientific processes as well as they might through slower-paced chalkboard explanations. We hope to address these issues in more depth in the next year by concentrating on laboratory sections and student interaction for a few selected cases.

Campus infrastructure, instructional support, and technical support are essential concerns of the faculty studied. They are willing, if not always eager, to experiment with computer-based technologies in the classroom, provided sufficient support is available. They want support for developing instructional materials such as I-scapes. The ADEPT project provided this support by investing a considerable amount of graduate research assistant effort. Similarly, they need technical support so that precious minutes in the classroom are not wasted with set-up, debugging, and take-down of equipment. The faculty we are studying all are sophisticated users (and some are developers) of information technologies. They know from experience that overhead projectors and chalk are more reliable instructional technologies than are computer systems, and these experiences are reflected in their teaching styles and their advice to the project.

Digital libraries hold great potential for teaching and learning at the undergraduate level. The ADEPT project is building on a rich source of geographic information resources in the Alexandria Digital Library, advanced technical infrastructures at two major research universities, and the participation of technically sophisticated faculty who teach undergraduate courses. We have a tremendous opportunity to understand more about the requirements for constructing digital libraries that will enhance scientific thinking and learning. We are making inroads at understanding the problem and hope to offer some workable solutions in later phases of the project.

Acknowledgments

This article is based on work supported by the National Science Foundation under grant no. IIS-9817432, Terence R. Smith, University of California, Santa Barbara, Principal Investigator. Our thanks go to our colleagues at UCSB and to the faculty and students at UCLA and UCSB who are participating in the ADEPT project. The article benefitted greatly from comments on earlier drafts by Michael Goodchild, UCSB; Mary Marlino, Digital Library for Earth Sciences Education; and Anita Coleman, UCLA.

References

Baxter, G. P.; Shavelson, R. J.; Goldman, S. R.; & Pine, J. (1992). Evaluation of procedure-based scoring for hands-on science assessment. *Journal of Educational Measurement, 29*(1), 1-17.

Bishop, A. P.; Buttenfield, B. P.; & Van House, N. (Eds.). (in press). *Digital library use: Social practice in design and evaluation.* Cambridge, MA: MIT Press.

Borgman, C. L. (in press-a). Digital libraries and virtual universities. In T. Della Senta & T. Tschang (Eds.), *Access to knowledge: New information technologies and the emergence of the virtual university.* Paris: International Association of Universities, and Oxford, England: Pergamon.

Borgman, C. L. (In press-b). Uses, users, and usability of digital libraries. In A. P. Bishop, B. P. Buttenfield, & N. Van House (Eds.), *Digital library use: Social practice in design and evaluation.* Cambridge, MA: MIT Press.

Borgman, C. L. (2000). *From Gutenberg to the global information infrastructure: Access to information in the networked world.* Cambridge, MA: MIT Press.

Borgman, C. L.; Bates, M. J.; Cloonan, M. V.; Efthimiadis, E. N.; Gilliland-Swetland, A.; Kafai,Y.; Leazer, G. L.; & Maddox, A. (1996). *Social aspects of digital libraries* (Final Report to the National Science Foundation; Computer, Information Science, and Engineering Directorate; Division of Information, Robotics, and Intelligent Systems; Information Technology and Organizations Program, 1996. Award number 95-28808). Retrieved September 8, 2000 from the World Wide Web: http://dlis.gseis.ucla.edu/DL/UCLA_DL_Report.html.

Buttenfield, B. P., & Kumler, M. P. (1996). *Tools for browsing environmental data: The Alexandria Digital Library interface* (Proceedings of the Third International Conference on Integrating Geographic Information Systems and Environmental Modeling, Santa Fe, NM, January 21-25, 1996). Santa Barbara, CA: National Center for Geographic Information and Analysis. Retrieved September 9, 2000 from the World Wide Web: http://www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/sf_papers/buttenfield_babs/babs_paper.html.

Council on Library and Information Resources. (1999). Executive summary. In *Scholarship, instruction, and libraries at the turn of the century: Results from five task forces appointed by the American Council of Learned Societies and the Council on Library and Information Resources.* Washington, DC: CLIR. Retrieved September 9, 2000 from the World Wide Web: http://www.clir.org/pubs/reports/pub78/preface.html#executive.

Doran, R. L.; Lawrenz, F.; & Helgeson, S. (1994). Research on assessment in science. In D. L. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 388-442). New York: Macmillan.

Fitz-Gibbon, C. T., & Morris, L. L. (1987). *How to design a program evaluation,* 2d ed. Newbury Park, CA: Sage.

Geography Education Standards Project. (1994). *Geography for life: National geography standards.* Washington, DC: National Geographic Research & Exploration.

Gilliland-Swetland, A. J. (1998). Evaluation design for large-scale, collaborative online archives: Interim report of the Online Archive of California Evaluation Project. *Archives & Museum Informatics, 12*(3-4), 177-203.

Hill, L. L.; Carver, L.; Larsgaard, M.; Dolin, R.; Smith, T. R.; Frew, J.; & Rae, M. A. (2000). Alexandria Digital Library: User evaluation studies and system design. *Journal of the American Society for Information Science, 51*(3), 246-259.

Kafai, Y., & Gilliland-Swetland, A. J. (in press). Integrating historical source materials into elementary science classroom activities. *Science Education.*

Leazer, G. H.; Gilliland-Swetland, A. J.; & Borgman, C. L. (2000). Evaluating the use of a geographic digital library in undergraduate classrooms: The Alexandria Digital Earth Prototype (ADEPT). In *Proceedings of the Fifth ACM Conference on Digital Libraries* (San Antonio, TX, June 2-7, 2000) (pp. 248-249). New York: ACM Press.

Leazer, G. H.; Gilliland-Swetland, A. J.; Borgman, C. L.; & Mayer, R. (in press). *Classroom evaluation of the Alexandria Digital Earth Prototype (ADEPT).* Unpublished paper presented at the American Society for Information Science, 2000 Annual Conference, Chicago.

Marchionini, G., & Crane, G. (1994). Evaluating hypermedia and learning: Methods and results from the Perseus Project. *ACM Transactions on Information Systems, 12*(1), 5-34.

Metz, K. (1995). Reassessment of developmental constraints on children's science instruction. *Review of Educational Research, 65*(2), 93-127.

Mose, D., & Maney, T. (1993). An experiment in distance learning of geology. *Journal of Computers in Mathematics and Science Teaching, 12*(1), 5-18.

National Research Council, Center for Science, Mathematics, and Engineering Education. (1998). *Developing a digital national library for undergraduate science, mathematics, engineering and technology education: Report of a workshop.* Washington, DC: National Academy Press.

National Research Council. (1996). *National science education standards: Observe, interact, change, learn.* Washington, DC: National Academy Press.

National Science Foundation. (1999). *Digital Libraries Initiative, Phase II.* Retrieved September 9, 2000 from the World Wide Web: http://www.dli2.nsf.gov.

Newnham, R.; Mather, A.; Grattan, J.; Holmes, A.; & Gardner, A. (1998). An evaluation of the use of Internet sources as a basis for geography coursework. *Journal of Geography in Higher Education, 22*(1), 19-34.

Nielsen, J. (1993). *Usability engineering.* Boston: Academic Press.

Persky, H. R.; Reese, C. M.; O'Sullivan, C. Y.; Lazer, S.; Moore, J.; & Shakrani, S. (1996). *NAEP 1994 Geography Report Card: Findings from the National Assessment of Educational Progress* (NCES 96-087). Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human-computer interaction,* 3d ed. Reading, MA: Addison Wesley Longman.

Shneiderman, B.; Byrd, D.; & Croft, W. B. (1997). Clarifying search: A user-interface framework for text searches. *D-Lib Magazine,* 3(1). Retrieved September 9, 2000 from the World WideWeb: http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html.

Soloway, E.; Norris, C.; Blumenfeld, P.; Fishman, B.; Krajcik, J.; & Marx, R. (2000). K-12 and the Internet. *Communications of the ACM, 43*(1), 19-23.

Toward an Every-Citizen Interface to the Nation's Information Infrastructure Steering Committee. Computer Science and Telecommunications Board; Commission on Physical Sciences, Mathematics, and Applications; National Research Council. (1997). *More than screen deep: Toward every-citizen interfaces to the nation's information infrastructure.* Washington, DC: National Academy Press.

Wallace, R.; Krajcik, J.; & Soloway, E. (1996). Digital libraries in the science classroom: An opportunity for inquiry. *D-Lib Magazine,* September 1996. Retrieved September 9, 2000 from the World Wide Web: http://www.dlib.org/dlib/september96/umdl/09wallace.html.